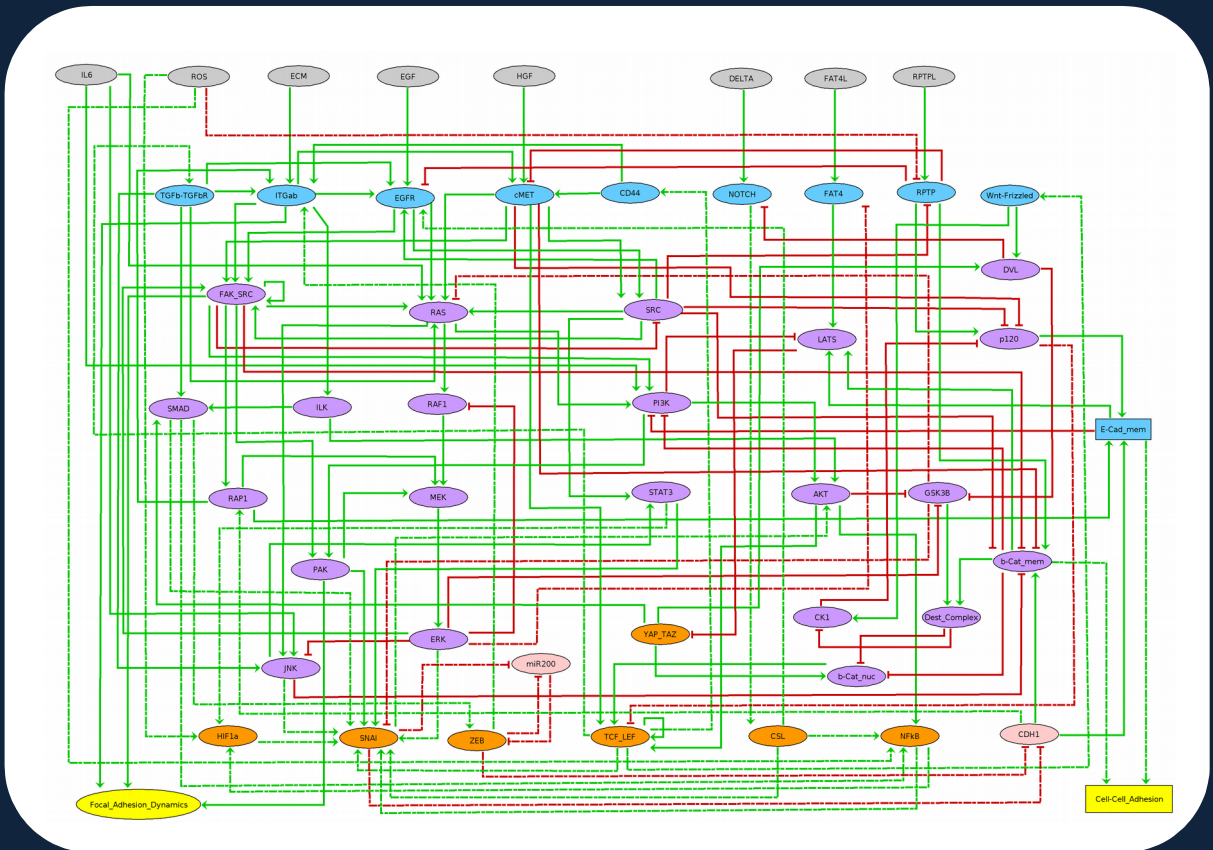


# How cells initiate Epithelial-to-Mesenchymal Transition?

A computational modelling of cellular and supra-cellular networks to unravel the control of EMT

Ricardo Jorge Fonseca Tavares Godinho Pais



Dissertation presented to obtain the Ph.D degree in Systems Biology  
Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Oeiras,  
September, 2018



UNIVERSIDADE  
**NOVA**  
DE LISBOA

# How cells initiate Epithelial-to-Mesenchymal Transition

A computational modelling of cellular and supra-cellular  
networks to unravel the control of EMT

Ricardo Jorge Fonseca Tavares Godinho Pais

Dissertation presented to obtain the Ph.D degree in Systems Biology  
Instituto de Tecnologia Química e Biológica António Xavier | Universidade Nova de Lisboa

Research work coordinated by:  
Dr. Claudine Chaouiya



Oeiras, September, 2018



## **Supervisor**

Dr. Claudine Chaouiya, principal Investigator of the Network Modelling Group of the Instituto Gulbenkian Ciência, Oeiras, Portugal.

## **Thesis committee**

Dr. Florence Janody, is the Coordinator of the CYTOSKELETAL REGULATION and CANCER group at IPATIMUP/i3S (Instituto de Patologia e Imunologia Molecular da Universidade do Porto/Instituto de Investigação e Inovação em Saúde) in Porto. Also an associated Principal Investigator at Instituto Gulbenkian de Ciencia (IGC), Oeiras, Portugal.

Dr. Jorge Carneiro, Principal Investigator of the Quantitative Organism Biology Group of the Instituto Gulbenkian Ciência, Oeiras, Portugal.

## **Funding**

This thesis was sponsored by Fundação para a Ciência e Tecnologia fellowship SFRH/BD/52175/2013 and Instituto Gulbenkian de Ciência. The last year was sponsored by BioenhancerSystems, a private company that develops analytical solutions and services.



## **Communications in meetings or international conferences**

Ricardo J. Pais & C. Chaouiya, "Computational modelling of the regulation of cell adhesion, critical step of EMT-like program", International conference on systems biology, Barcelona, September 2016. DOI:[10.6084/m9.figshare.5501032.v2](https://doi.org/10.6084/m9.figshare.5501032.v2).

Ricardo J. Pais & C. Chaouiya, "Computational modelling of the regulation of cell adhesion, critical step of EMT-like program, 10th Annual Meeting of Gulbenkian Students (AmeeGuS), Setúbal, 17-19 May 2016. DOI:[10.6084/m9.figshare.6983336](https://doi.org/10.6084/m9.figshare.6983336).

Ricardo J. Pais & C. Chaouiya, "A Computational modelling of the regulation of cell adhesion fate, first step of EMT-like program", 9th Annual Meeting of Gulbenkian Students (AmeeGuS), Mondim de Basto, Portugal, 17-21 of May 2015. DOI:[10.6084/m9.figshare.5500906.v3](https://doi.org/10.6084/m9.figshare.5500906.v3).

Ricardo J. Pais & C. Chaouiya, "A Computational modelling of the regulation of cell adhesion fate, first step of EMT-like program", XXIII Porto Cancer Meeting, Porto, 7-8 of May 2015. DOI:[10.6084/m9.figshare.5500906.v3](https://doi.org/10.6084/m9.figshare.5500906.v3).



## **Acknowledgements**

I would like to thank all the people that directly or indirectly contributed to this thesis. Innumerous people have contributed either with motivational support or with scientific knowledge through discussions and opinions. However, I would like to highlight and thank here the following people that had a more direct role in making this thesis work possible:

- My supervisor Claudine for accept me in her group and for giving me the opportunity to address the Epithelial-to-Mesenchymal transition in the cancer context as my Ph.D thesis project. I would also like to thank Claudine for the opportunity of learning the logical modelling framework and apply it to the regulation of biological systems.
- The IGC Ph.D Program director Élio Sucena for believing in my capacity and value as a scientist since the day of the IGC interview. I am truly thankful for all his pragmatic advices and huge efforts to ensure the best conditions possible for all program candidates. Also for all efforts conducted in helping through the process of making a PhD in IGC.
- The IGC Network Modelling Group for the good working environment and scientific discussions during the time of Ph.D. In particular, José Curi, for all wonderful scientific discussions, wisdom and friendship that keep me going further.
- A very special thanks to Pedro Fernandes, the coordinator of the Gulbenkian Training Programme in Bioinformatics (GTPB) for the friendship, wise advices, external and independent guidance, amazing courses and frequent discussions in the last five years. In addition, I would also like to credit the GTPB courses that I attended, specially the "Programming in Python for biomedical researchers" course, because they provided me the tools and skills to solve several problems in the work presented in this thesis.
- To all amazing friends, colleges and brilliant researchers that I have meet in the IGC during the time of the Ph.D (they are too many to mention).

Among all, I would like to specially thank Elves Duarte, José Santos, Lok Pahari, Abdulbasit Amin, and Carlos Alexandre Elias for the brilliant discussions, and for standing always by my side even in the bad moments, true friends.

Unfortunately, referring all people is not possible since it would give another thesis. If you are such person and not mentioned here, doesn't mean that I did not remember you. For all those special people a very big and warm thank you! You have always a place in my heart.

Quero aqui também agradecer à minha linda companheira Sandra Bráz pelo seu amor e apoio incondicional durante todas as adversidades e sacrifícios que foram feitos ao longo dos últimos 3 anos. Quero também agradecer à minha avó Julieta Pais e tia Ana Paula Pais por todo apoio moral e as ajudas que me deram comigo e com o Tiago ao longo destes últimos 5 anos.

Por fim, quero agradecer ao meu filho Tiago Pais, pela força que me deu, e pela compreensão que demonstrou apesar da sua tenra idade. Tiago, és a pessoa mais importante da minha vida e todos os esforços que fiz e feitos que consegui na minha vida foram e serão sempre dedicados a ti.

**“Sem ti Tiago, meu filho e companheiro de armas, nunca teria a força sobrehumana para chegar aqui, foste a minha fonte regeneradora e motivação desde 2006”.**

**“Sem ti Sandra não seria a mesma coisa, não teria o mesmo sabor, muito obrigado por tudo mulher linda”.**

## **Thesis outline**

Over the last decade, Epithelial-to-Mesenchymal Transition (EMT) has gained the attention of cancer researchers due to its potential to promote cancer migration and metastasis. However, the complexity of EMT intertwined regulation and the involvement of multiple signals in the tumour microenvironment have been limiting the understanding of how this process can be controlled. Cell-cell adhesion and focal adhesion dynamics are two critical properties that change during EMT, which provide a simple way to characterize distinct modes of cancer migration. Therefore, the main focus of this thesis is to provide a framework to predict critical microenvironment and de-regulations in cancer that drive inter-conversion between adhesion phenotypes, accounting for main microenvironment signals and signalling pathways in EMT. Here, we address this issue through a systems approach using the logical modelling framework to generate new testable predictions for the field.

In Chapter 1, we describe the basic features that characterize the process of Epithelial-to-Mesenchymal Transition, including changes in cell adhesion and its role in the process of metastasis of carcinomas. Here, we highlight involved signalling pathways, main reported microenvironment signals and key transcription factors that regulate EMT. In this chapter, we focus on E-cadherin mediated cell-cell adhesion and integrin mediated focal adhesion dynamics as key cell adhesion properties that characterize Epithelial, Mesenchymal, and partial EMT phenotypes. We review the literature for the existent mathematical models of EMT and highlight their main focus and predictions. Finally, we provide a description of the main concepts of the logical modelling framework and highlighted the reasons to choose this approach.

In Chapter 2, we describe the developed model of the regulation of cell-cell adhesion and the focal adhesion dynamics, taking 8 microenvironment signals as inputs and the 2 cell adhesion properties as readouts. In this chapter, we validate the model by recapitulating phenotypic markers and observations on cell adhesion changes from published experiments on Epithelial cell lines exposed to several microenvironment signals or containing mutations on specific molecular

components. In addition, we also recapitulate activity changes in model components described in the above mentioned experiments. In this chapter, we associate cell adhesion properties to phenotypes and analyse the conditions that result in the switching between phenotypes. Strikingly, our model analyses showed that cell-cell contact signals activating Receptor-Protein Tyrosine Phosphatases (RPTP) are key microenvironment signals to control EMT, promoting the Epithelial phenotype. On the other hand, we predict that the ectopic activation of another transmembrane protein regulated by cell-cell contact signals (FAT4) can promote the stability of collective migration phenotypes. This protein is observed to be overexpressed in several cancers supporting the idea that FAT4 over-activation could play a role in cancer invasion. Moreover, with our approach, we also provide an explanation for the observed correlation between hypoxia and metastasis, reinforcing the idea that uncontrolled SRC ectopic activation is enough to promote the stability of single cell cancer migration phenotypes.

In Chapter 3, the main predictions generated by our model and inherent limitations are discussed. In this chapter, we also outline possible future work and proposed some experiments to test our predictions.

In summary, this thesis provides a new mathematical model and a workflow to generate predictions that explore the combinations of microenvironment and mutational effects on the switching between cell adhesion properties. Importantly, this thesis contributes to the field with the identification of a not yet reported mechanism of controlling EMT and propose new hypotheses that explain the stability of collective and single cell migration of carcinomas.

### **Resumo da Dissertação**

Na última década, a transição Epitelial para Mesenquimal (EMT) tem chamado a atenção dos cientistas derivado ao seu potencial para promover a migração de células cancerosas e metastases. Contudo, a complexidade da regulação da EMT, envolvendo múltiplos sinais do microambiente tumoral e regulação cruzada tem sido uma limitação para perceber como se pode controlar EMT em tumores.

A adesão célula-célula e a dinâmica de adesão focal são duas propriedades críticas que alteram durante a EMT e constituem uma maneira simples de caracterizar distintos modos de migração de células cancerosas. Sendo assim, o principal foco desta tese é proporcionar um modelo matemático para prever condições críticas que levam à transição entre fenótipos com diferentes tipos de adesão celular, tendo em conta os principais sinais do microambiente e desregulações frequentes em células cancerosas. Nesta tese, é proposto um modelo matemático integrativo baseado no formalismo lógico para gerar novas previsões para este campo.

No capítulo 1, apresentamos uma descrição das características principais do processo da EMT, incluindo as alterações nas propriedades adesivas das células e o seu papel no processo de metastases em carcinomas. Aqui, as principais vias de sinalização, sinais do microambiente e factores de transição envolvidos na EMT são sublinhados. Neste capítulo, a adesão célula-célula mediada pela E-caderina e a adesão focal mediada por integrinas como propriedades que caracterizam as células Epiteliais, Mesenquimais, e fenótipos resultantes de EMT parcial são apresentados em promenor. Também é apresentado aqui, uma revisão da literatura referente aos modelos matemáticos existentes em EMT, sublinhando os seus principais objectivos e previsões. Por fim, descrevemos aqui os principais conceitos da modelação lógica e apresentamos as razões para a escolha desta metodologia.

No capítulo 2, apresentamos o modelo lógico da rede intracellular da regulação das propriedades adesivas desenvolvido neste trabalho, considerando 8 sinais do microambiente como variáveis de entrada e 2 propriedades de adesão celular como variáveis de leitura. Neste capítulo, validamos o modelo comparando os seus resultados com marcadores fenotipicos e observações nas propriedades adesivas e actividade intracelular retiradas experiências publicadas em células Epiteliares expostas a sinais extracellulares ou contendo mutações em seus components moleculares. Os resultados do modelo aqui apresentados são compatíveis com todas as observações experimentais tanto a nível fenotipico como em termos de actividade experimental, suportando a sua validação. Neste

capítulo, as propriedades adesivas foram correspondidas com os fenótipos celulares e analisou-se as condições que levam à inter-conversão entre fenótipos. Surpreendentemente, a nossa análise mostrou que sinais que provêm de contactos célula-célula que activam as fosfatases de tirosinas membranares (RPTPs) são sinais do microambiente que têm um papel chave no controlo da EMT. Por outro lado, o modelo gerou previsões que sugerem os contactos célula-célula que resultam na activação da proteína membrana FAT4 pode levar a estabilização de fenótipos envolvidos na migração colectiva de células cancerosas. Esta proteína foi observada estar sobre-expressa em tumores, o que suporta a ideia que a sobre activação de FAT4 pode ter um papel na invasão de tumores. Por fim, a abordagem apresentada nesta tese permitiu uma explicação para a correlação entre hipóxia e metastases, e permitiu reforçar a ideia que a activação ectópica de SRC é suficiente para promover a estabilidade de fenótipos envolvidos na migração unicelular de células cancerosas.

No capítulo 3, as principais previsões do modelo e suas limitações são discutidas. Neste capítulo, também é referido possível trabalhos futuros de continuação e algumas experiências para testar as previsões do modelo são propostas.

Em suma, esta tese contribui com um novo modelo matemático e uma metodologia que permite gerar previsões que exploram sistematicamente as varias combinações de microambientes e efeitos mutacionais que levam a mudanças nas propriedades adesivas das células. Os resultados expressos nesta tese constituem um contributo relevante para a área com a identificação de um novo mecanismo de controlo da EMT e propõe novas hipóteses que explicam a estabilidade da migração individual e colectiva de carcinomas.

## **Abstract**

The objective of this thesis is to provide an integrative (holistic) computational approach to better understand how Epithelial-to-Mesenchymal transition (EMT) is controlled by the microenvironment in the context of cancer and metastasis. Since EMT regulation is complex, we focus on cell-cell adhesion and focal adhesion dynamics as proxies for the main phenotypes involved in the process of metastasis (Epithelial, Mesenchymal and Hybrid phenotypes). Therefore, we developed a literature-based logical network model, accounting for 8 key microenvironment signals as model inputs, 9 key signalling pathways or regulatory mechanisms, and two cell adhesion properties as model readouts that describe phenotypes. The model was able to recapitulate the observed activity degrees of 9 Epithelial/Mesenchymal markers, reported activity patterns of pathways, and 169 observations from 13 experiments on Epithelial cell lines. Next, we aimed to interrogate the model to generate novel predictions for the control of the switching between phenotypes. Strikingly, the model-based analysis showed that RPTP activation by cell-cell contacts (model input) prevented the switching from Epithelial to Mesenchymal/Hybrid phenotype in the presence of signals that promote EMT. The model-based analysis also showed that RPTP activation triggered the switching from Mesenchymal/Hybrid to Epithelial phenotype. These results suggest that RPTP could play a role as a control mechanism of EMT. In addition, the model-based analysis showed that the combination of ROS with an EMT promoting signal inhibited the RPTP control mechanism. This explained the observed correlation between metastasis and the combination of hypoxia with chronic inflammation signals. Perturbation analysis on model components (e.g KO or Ectopic activation) resulted in the identification of 23 single and 24 double perturbations that either stabilized the Hybrid phenotype or destabilized the Epithelial/Mesenchymal phenotypes, suggesting possible critical intra-cellular alterations. Among these, SRC ectopic activation destabilized the Epithelial phenotype and FAT4 ectopic activation stabilized the Hybrid phenotype. These perturbations were correlated with high overexpression in carcinomas supporting their role in carcinoma invasion.

Further, the model-based analysis showed that a perturbation that mimics the SRC overexpression prevents the RPTP control, explaining the correlation between this alteration and cancer invasion. Moreover, the model-based analysis showed that cell-cell contacts that activate FAT4 in the case of a perturbation that mimics FAT4 overexpression modulate the transitions between Hybrid and Mesenchymal phenotype. In conclusion from this thesis, RPTP and FAT4 activation by cell-cell contacts are key microenvironment signals that are capable of controlling the transitions between Epithelial, Mesenchymal and Hybrid phenotype.



## Table of Contents

CHAPTER 1: Introduction.....	8
1. Thesis Scope.....	9
2. Epithelial-to-Mesenchymal Transition (EMT).....	10
2.1. Physiological function.....	12
2.2. EMT functional steps.....	14
2.3. Partial EMT phenotypes.....	16
2.4. EMT in carcinomas.....	20
2.4.1. Modes of invasion and metastasis.....	21
2.4.2. Cell adhesion role in carcinoma invasion.....	23
2.4.3. The role of the microenvironment.....	27
2.5. Transcriptional regulation of EMT.....	29
2.6. Signalling pathways involved in EMT.....	31
3. Addressing EMT through a Systems Approach.....	35
3.1. The logical formalism.....	38
3.1.1. Advantages and limitations.....	40
3.1.2. Model definition.....	42
3.1.3. Model dynamics.....	44
3.1.4. Methods and tools.....	47
3.1.5. Modelling cancer networks.....	48
3.2. Mathematical models of EMT.....	51
3.2.1 Kinetic models of EMT.....	51
3.2.2. Logical models of EMT.....	55
3.3. Modelling the effects of microenvironment signals on cell adhesion.....	57
4. References.....	58
CHAPTER 2: Logical Modelling of the Regulation of Cell Adhesion properties in EMT.....	72
Summary.....	73
1. Introduction.....	74

2. Methods.....	77
2.1. Software tools.....	77
2.2. Model construction.....	78
2.3. Model static analysis.....	84
2.4. Analysis of model dynamics.....	86
2.5. Model validation data and comparisons.....	89
3. Results.....	93
3.1. Modelling workflow.....	93
3.2. Logical network model of cell adhesion properties.....	95
3.3. Recapitulation of model phenotypes.....	98
3.4. Recapitulated biological markers and patterns.....	100
3.5. Recapitulated experiments on Epithelial cell lines.....	102
3.6. Phenotypes compatibility with the microenvironment.....	105
3.7. The reachability of adhesion phenotypes.....	107
3.8. The effect of model perturbations on phenotypes.....	109
3.9. The impact of SRC and FAT4 overexpression.....	112
3.10. The mechanism of RPTP and FAT4 on cell adhesion.....	116
4. Discussion.....	119
5. Supplemental Information.....	126
5.1. Supplemental text.....	126
5.2. Supplemental figures.....	137
5.3. Supplemental tables.....	146
6. References.....	160
CHAPTER 3: Discussion & Conclusion.....	176
1. Discussion.....	177
1.1. A computational model of cell adhesion.....	177
1.2. The role of cell-cell dependent RPTPs.....	179
1.3. Critical deregulations promoting EMT in carcinomas.....	181
1.4. The stability of Hybrid phenotypes.....	182
2. Conclusions.....	183

3. Future work.....	184
4. References.....	185

## Figures and Tables

Figure 1. Typical Epithelial and Mesenchymal features involved in Epithelial-to-Mesenchymal Transition (EMT).....	11
Figure 2. EMT archetype and Epithelial plasticity.....	16
Figure 3. Types and variants of collective cell migration.....	19
Figure 4. Modes of invasion of carcinomas and metastasis acquisition.....	22
Figure 5. Critical cell adhesion changes during EMT progression.....	26
Figure 6 Roles and regulation of major EMT transcription factors.....	31
Figure 7. Regulatory mechanisms of TGF- $\beta$ induced EMT.....	34
Figure 8. Signalling pathways involved in EMT progression.....	35
Figure 9. The systems biology cycle.....	38
Figure 10. Modelling the time-course kinetics using a kinetic model and the correspondent logical abstraction.....	41
Figure 11. Formal logical model definition, an illustrative example.....	43
Figure 12. Model dynamics, an illustrative example.....	46
Figure 13. Network representation of the mechanisms of EMT by EGF and Wnt signalling, modelled in the work of Shin et al.....	52
Figure 14. The proposed core regulatory circuit of EMT.....	54
Figure 16. Examples of annotations in GINsim annotation box.....	80
Figure 17. Examples of biological mechanisms and their correspondent logical abstraction.....	82
Figure 18. Workflow for the logical network model construction, validation and analysis.....	94
Figure 19. Regulatory network of cell-cell adhesion and focal adhesion dynamics.....	96
Figure 20. Combinations of the model outputs (adhesion properties) and compatible phenotypes.....	99
Figure 21. Representation of the model stable states of cell adhesion properties of Epithelial (A), Mesenchymal (B), and intermediate (C) phenotypes.....	101
Figure 22. Comparison of the observations of experiments on Epithelial cell lines (top) with simulation results (bottom).....	104

Figure 23. The number of microenvironments compatible with adhesion phenotypes for different combinations of RPTP and FAT4 ligands.....	106
Figure 24. Reachability probabilities of cell adhesion phenotypes in physiologic scenarios. ....	108
Figure 25. Effects of single and double perturbations on the model adhesion phenotypes. ....	112
Figure 26. Comparison of the model phenotype switching behaviours due to cell-cell contact signals with and without perturbations that mimic SRC and FAT4 overexpression. ....	114
Figure 27. Critical interactions that are involved in the control the cell adhesion properties under chronic inflammation conditions with RPTPL (A) and FAT4L (B) signals.....	117
Figure S1. Adherens junctions regulatory map of KEGG PATHWAYS database (map hsa04520).....	137
Figure S2. Focal adhesion regulatory map of KEGG PATHWAYS database (map hsa04510).....	138
Figure S3 Hippo signalling map of KEGG PATHWAYS database (map hsa04590).....	139
Figure S4 . NOTCH signalling map of KEGG PATHWAYS database (map hsa04530)..	140
Figure S5. High cell-cell adhesion and basal focal adhesion dynamics (C=2 & F= 0) stable states.....	141
Figure S6. Basal cell-cell adhesion and high focal adhesion dynamics (C=2 & F= 0) stable states.....	141
Figure S7. Intermediate cell-cell adhesion and basal focal adhesion dynamics (C=1 & F= 0) stable states.....	142
Figure S8. Number of observations in the experiments on epithelial cell lines used for model validation.....	142
Figure S9. Comparison of model stable states for Epithelial and Mesenchymal phenotypes with the ones of a published EMT model.....	143
Figure S10. Simulations of the unperturbed model for normal A-B and chronic inflammation conditions C-D.....	144
Figure S11. Simulations of the model with SRC perturbation (SRC[FAK_SRC@0]) for normal A-B and chronic inflammation conditions C-D.....	144
Figure S12. Simulations of the model with FAT4 perturbation (FAT4[ERK@0]) for normal A-C and chronic inflammation conditions D-F.....	145

Table 1. Cell adhesion properties of invasive and non-invasive phenotypes.....	26
Table 2. Summary of logical modelling software tools.....	49
Table 3 . KEGG reference maps used for network construction.....	79
Table 4. Controlled vocabulary used for the description of biological mechanisms and their correspondent time scales.....	81
Table 4. EMT biomarkers used for model validation.....	89
Table 5. Selected experiments on mammalian epithelial cell lines used for model validation, grouped in perturbation type.....	90
Table S1. Network model components.....	146
Table S2. Interactions of the network and their effects.....	148
Table S3. Model logical functions and the correspondent type of regulatory mechanisms (R.M.).....	152
Table S4. Parameters used in the simulations of experiments on epithelial cell lines.....	154
Table S5. Nodes of the Steinway EMT logical network model not included in the logical network model of the regulation of cell adhesion properties.....	154
Table S6. Model input conditions (microenvironment conditions) compatible with the model adhesion phenotypes for each combination of cell-cell contact signals (RPTP and FAT4 ligands).....	156
Table S7. Model input configurations compatible with physiological scenarios used in simulations.....	156
Table S8. Comparison of the effects of single perturbations on model components with the expression of these proteins in human carcinomas.....	157
Table S9. Comparison of the effects of double perturbations on model components with the expression of these proteins in human carcinomas.....	158
Table S10. Impact of perturbations on model interactions in terms of the reachability of adhesion phenotypes under chronic inflammation and RPTPL signals.....	159

## List of Abbreviations

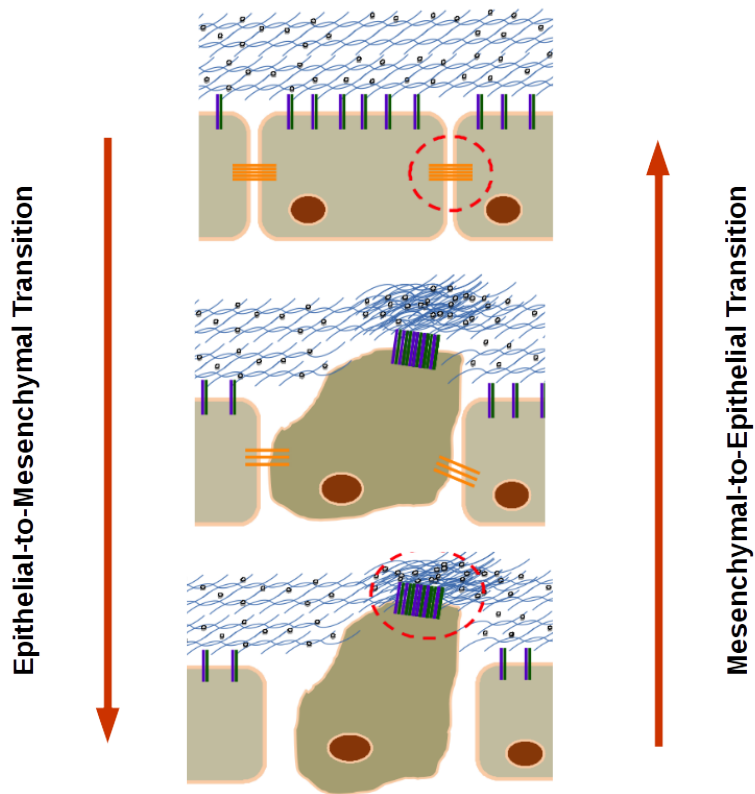
ACSN	Atlas of Cancer Signalling Network
AKT	RAC- $\alpha$ serine/threonine-protein kinase
CK1	Casein kinase 1
cMET	Hepatocyte growth factor receptor
DNA	Deoxyribonucleic acid
DVL	Human homologue of segment polarity protein dishevelled
ECM	Extra-cellular matrix
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
ELISA	Enzyme-Linked Immunosorbent Assay
EMT	Epithelial-to-Mesenchymal Transition
FAK	Focal Adhesion Kinase
FAT4L	FAT4 protein Ligand
GSK3 $\beta$	Glycogen synthase kinase 3 beta
H1975	Non-small lung cancer cell line from human adenocarcinoma
HET	Hybrid-to-Epithelial transition
HGF	Hepatocyte Growth Factor
HIF1	Hypoxia Inducing Factor 1
HTG	Hierarchical Transition Graph
IL6	Interleukin 6
ILK	Integrin-linked protein kinase
JAK	Janus kinase
JNK	c-Jun N-terminal kinase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LATS	Large tumour suppressor complex
LEF	Lymphoid enhancer-binding factor 1
MAPK	Mitogen-Activated Protein Kinase
MAT	Mesenchymal-to-Amoeboid Transition
MCF10A	Human mammary epithelial cell line
MDCK	Madin-Darby Canine Kidney cell line
MET	Mesenchymal-to-Epithelial Transition
miR200	Micro RNA 200

MMP	Matrix metalloproteinase
NFkB	Nuclear factor NF-kappa-B complex
ODE	Ordinary Differential Equations
p120	Catenin delta
PAK	Serine/threonine-protein kinase PAK
PI3K	Phosphatidylinositol 4,5-bisphosphate 3-kinase
qPCR	Quantitative Polymerase Chain Reaction
ROS	Reactive Oxygen Species
RPTP	Receptor-type Protein Tyrosine Phosphatase
RPTPL	Receptor-type Protein Tyrosine Phosphatase Ligand
RTK	Receptor Tyrosine Kinases
SHH	Sonic hedgehog
SMAD	Mothers against decapentaplegic homologue protein
Snail	Human homologue Zinc finger protein
SRC	Proto-oncogene tyrosine-protein kinase Src
STAT	Signal transducer and activator of transcription
STG	State Transition Graph
TAZ	WW domain-containing transcription regulator protein 1
TCF	T-cell-specific transcription factor
TGF-β	Transforming Growth Factor beta
TGF-βR	Transforming Growth Factor beta receptor
Twist	Twist-related protein
YAP	Yes-associated protein 1
Zeb	Zinc finger E-box-binding homeobox protein

# CHAPTER 1

---

## Introduction





## 1. Thesis Scope

The main goal of this thesis is to provide an integrative (holistic) approach to better understand how carcinoma cells gain the capacity to migrate and to colonize other organs through Epithelial-to-Mesenchymal transition (EMT). In particular, to predict critical conditions that drive the inter-change between non-migrating and distinct forms of cell migration. This has been a hard challenge in cancer research due to the complexity of EMT regulation and by the involvement of multiple signals in the tumour microenvironment. To achieve our goal, we first aim to develop and analyse a logical network model for the regulation of cell adhesion properties involved in EMT and cancer invasion, accounting for multiple microenvironment signals and de-regulations that may occur during cancer progression. This provides a simple way to characterize distinct modes of cancer migration based on their adhesion properties. With this approach, we aim to generate new predictions that allow the identification of the conditions that promote the transitions between non-invasive Epithelial cell adhesion and invasive Mesenchymal/Hybrid cell adhesion. In particular, the tumour microenvironment conditions that are critical for the switching between non-invasive and invasive phenotypes, accounting for the not yet explored neighbouring cell contact signals in combination with signals from the ECM. Secondly, we aimed to understand how mutations during cancer progression could affect the dynamics of the transitions between phenotypes towards favouring cancer invasion. In our opinion, the resulting predictions from this study are important clues for the identification of new drug targets and the development of new therapeutical strategies to control metastasis.

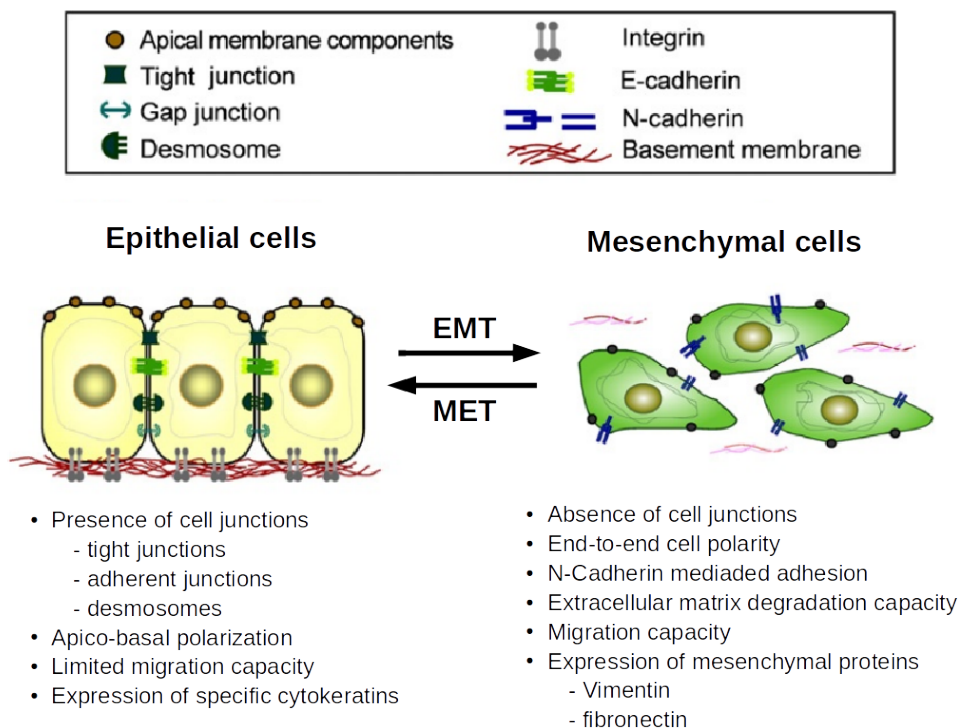
In this chapter, we provide the background on the process EMT focusing on the E-cadherin and integrin mediated cell adhesion changes and its role in the process of metastasis of carcinomas (section 2). We also included a literature

review on the existent mathematical models of EMT and provide a description of the main concepts of the logical modelling framework (section 3).

## **2. Epithelial-to-Mesenchymal Transition (EMT)**

Before defining EMT, we need to define what is an Epithelial cell, the Epithelium and the Mesenchymal cell. An Epithelial cell is a type of cell that is the “building blocks” of the membrane layer that surrounds most organs and glands [1,2]. In these tissues, Epithelial cells form sheets of cells attached by strong cohesive adhesion forces, in which the population of these cells is designated by Epithelium [1,3]. The main features that define Epithelial cells are the presence of distinct cellular structures called adherens junctions, tight junctions, desmosomes and gap junctions (Figure 1) [3]. These structures participate in the acquisition of the strong cell-cell adhesion between the cell and its adjacent cells. In addition, the spacial organization of the adhesion molecules on cell junctions together with the actin cytoskeleton determine the characteristic apico-basal cell polarity of Epithelial cells (Figure 1) [3]. Morphologically, Epithelial cells can adopt several shapes which can provide distinct tissue structure and functionalities, where the most common are the squamous, cuboidal and columnar [1]. On the other hand, Mesenchymal cells are a multi-potent type of stromal cells that can be found between endoderm and ectoderm during embryogenesis [4,5]. Mesenchymal cells do not show the presence cell adhesion structures mentioned above, and express characteristic proteins such as N-Cadherin, Vimentin and Fibronectin [3,4,6]. Mesenchymal cells are characterized by having elongated spindle shape morphology with end-to-end cell polarity, and the capacity of migrating along large distances either individually or in chains through N-Cadherin mediated cell adhesion (Figure 1) [2–4]. Moreover, this type of cells have also the capacity of

degrading and remodelling the extracellular matrix by secreting matrix metalloproteinases (MMP), collagen and fibronectin [4,7].



**Figure 1. Typical Epithelial and Mesenchymal features involved in Epithelial-to-Mesenchymal Transition (EMT).** Figure (adapted) and legend reproduced from [3].

By definition, EMT is a process in which an Epithelial cell undergoes multiple biochemical changes leading to the loss of the apico-basal cell polarity and cell-cell adhesion, followed by a “transformation” into a Mesenchymal cell with migratory capacity (Figure 1) [4,5,8,9]. Interestingly, EMT can be reverted through a process called Mesenchymal-to-Epithelial transition (MET) [4,8,10]. Both EMT and MET are classified as transdifferentiation processes based on their ability to occur between fully mature differentiated Epithelial and Mesenchymal cell types [10,11]. EMT is the result of a complex genetic program that depends on the

tissue context, involving multiple genes transcribed in Epithelial and Mesenchymal cells [4,5,8]. Multiple markers for EMT have been proposed during the last two decades [4,12]. However, EMT cellular program can be incomplete with strong Epithelial plasticity keeping both Epithelial and Mesenchymal features [5,8,13]. Due to such plasticity and complexity of EMT program, researchers have proposed the notion of EMT-like and define EMT based on functional criteria of phenotype changes [8,13].

## **2.1. Physiological function**

In the last two decades, EMT has been studied by researchers and observed as a process that plays a role in several physiological and pathological conditions, including embryogenesis, wound healing, organ fibrosis and cancer [3–5]. Historically, EMT was first observed in gastrulation and neural crest cell migration during early embryo development [6,9,14]. Since then, EMT has been observed in other developmental processes such as placental implantation and during the morphogenesis of many tissues in early embryogenesis. In general, EMT together with its reverse (MET), are two key steps that mediate the formation of the endoderm and mesoderm layers during organ development [4,9,14]. These steps are essential for the establishment of the epithelial barrier that delimits organs [4,9,14]. For example, through the occurrence of EMT, the primitive epithelium gives rise to the primary mesenchyme layer [4,9,14]. Then, the Mesenchymal cells migrate and through MET originates the secondary epithelial layer [4,9,14]. Several *in vivo* developmental studies in *Drosophila*, *Xenopus* and mouse models have documented the functional role of EMT during embryogenesis [6]. The outcomes from these studies have placed EMT as an evolutionarily conserved mechanism, essential for all multicellular organisms that develop organs [9].

After embryogenesis, EMT is also a necessary process to ensure the maintenance of epithelial tissue integrity of multicellular organisms during acute infections and injuries [2,3]. Mouse and *Xenopus* studies have provided insights on the role of EMT in epithelial tissue regeneration during wound healing and inflammatory damage [2–4]. Upon injuries, massive cell death occurs in the epithelium followed by a re-epithelialization of the tissue [3]. In response to damage, fibroblasts and inflammatory cells are recruited and secrete growth factors, cytokines and collagen type I [2–4]. Under these conditions, EMT program in functional Epithelial cells is naturally reactivated [3]. Then, the transformed Mesenchymal cells migrate to the damaged region and promote the decomposition of the basal membrane, causing apoptosis of damaged Epithelial cells [3]. Once the inflammatory signal is gone, the Mesenchymal cells undergo MET and re-epithelize the damaged epithelium [2,3]. Therefore, EMT in wound healing and acute inflammation is characterized by a transient and self-limiting process [3]. This transient behaviour is considered as normal, whereas the permanent stimulation of EMT due to chronic inflammation is pathological [3]. Consequently, chronic inflammation further results in tissue fibrosis, excessive cell death due to the accumulation of Mesenchymal cells, which eventually leads to organ failure or cancer development [3].

In the cancer context, EMT and its reverse (MET) captured the attention of researchers due to its potential to constitute a way for cancer cells to invade from the organ where they developed, migrate and metastasize [4,5,15]. This motivated researchers to study EMT in several types of cancer to find key players which could be used as drug targets to prevent metastasis [4,5,7]. However, the role of EMT in promoting metastasis is still a debatable question since some evidences points towards not being necessary, together with supporting evidences for its role only based on correlation [11,15]. Meanwhile, these

evidences are both weak and can be case/tissue specific due to the diversity of cancer genetic alterations together with the complexity and Epithelial plasticity of the EMT cellular program.

## **2.2. EMT functional steps**

Because EMT is a complex cellular program, researchers have used functional criteria to define EMT [8,13]. Based on developmental studies and according to Savagner [6], EMT can be defined and broken down into the following four independent functional steps (Figure 2).

**Step 1: Loss of apico-basal cell polarity with cell-matrix adhesion and cytoskeleton remodelling** [6]. This is considered as the initial step of EMT and characterized by the dissolution of tight junctions and desmosome structures, leading to a homogeneous distribution of apical and basolateral membrane components [5,6,8]. On this step, multiple membrane proteins and complexes that compose tight junctions are down-regulated, including the EMT marker ZO-1 [5,10]. Here, specific integrins that participate in cell polarity are down-regulated, including the ones that mediate cell to cell junctions and the ones that mediate cell to basement membrane connections [5,6]. During this step, the cytoskeleton is remodelled, where cytokeratin is replaced by vimentin and F-actin synthesis is increased (an G-actin polymer also named by stress fibres) [5,10]. Together these molecular changes result in a change from Epithelial cell shape (e.g. cuboidal) to a Mesenchymal shape (spindle) [5,6,8].

**Step 2: The loss of cell-cell adhesion forces and cell individualization** [6]. This step results in a phenotype change which is characterized by the dissolution of the adherens junction [6,16]. On the adherens junctions, catenins  $\alpha$ ,  $\beta$  and  $\delta$  are a class of proteins that directly interact with E-cadherin and subsequently participate in ensuring cell-cell adhesion strength [5,10,16]. E-cadherin is a key

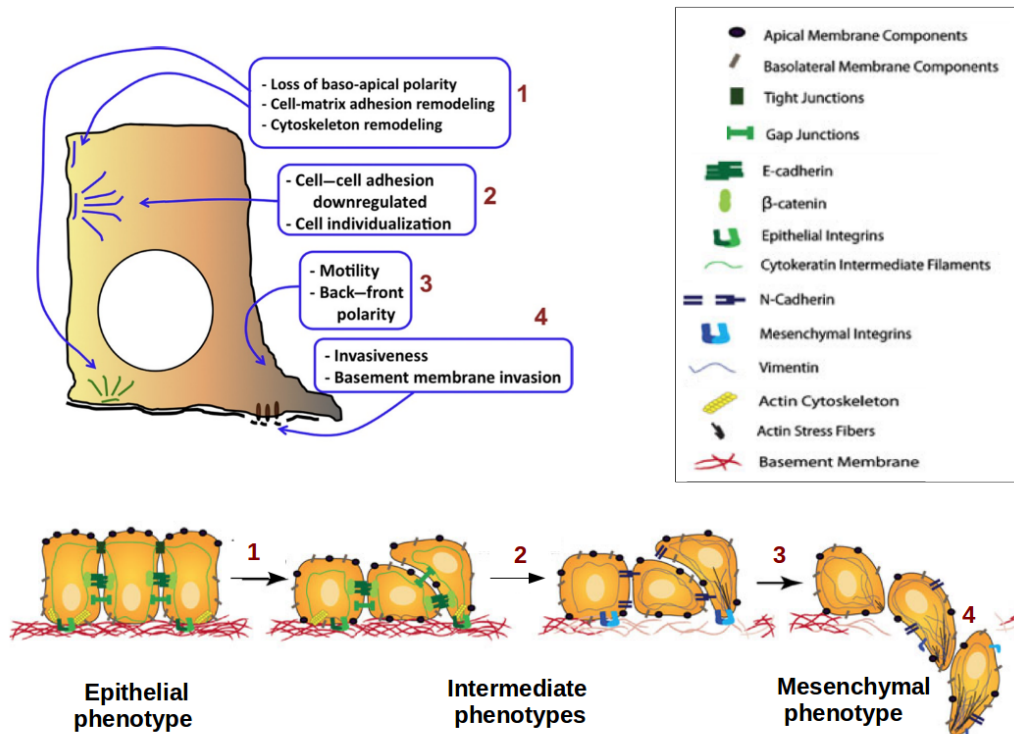
adhesion molecule of the adherens junction and its down-regulation is often seen as the hallmark of EMT, leading to cell individualization [4,5,10,12]. The loss of E-cadherin expression is subsequently replaced by N-cadherin, a less adherent molecule, which is a Mesenchymal marker [5,8,10].

**Step 3: Acquisition of back-front polarity and motility** [6]. During this step, cells acquire the capacity to migrate along the extracellular matrix (ECM) [5,6,10]. This is characterized by the formation of long sheet-like actin-rich membrane protrusions (lamellipodia) and spike-like extensions (filopodia) at one side of the cell (front) [5,6,10]. These membrane protrusions penetrate and temporarily interact with the ECM through the formation of localized cell-to-matrix adhesions (focal adhesions) [5,6]. Movement is then achieved through a front-rear pulling force generated by a mechanism of assembling and disassembling of focal adhesions, causing contraction and generating lamellipodia, filopodia, or invadopodia [5,6,10]. This step requires the activation of several proteins including small Rho GTPases, myosin kinases, and the expression of  $\alpha\beta1$  integrins [5,10].

**Step 4: Basement membrane invasion and invasiveness acquisition** [6]. During this final step, cells gain the capacity to decompose the basement membrane components by secreting matrix-metalloproteases (MMPs) [5,8,10,17]. In addition, cells also produce and secrete ECM components such as collagen type I and fibronectin that remodel the ECM to facilitate migration and invasion [4,5,18]. The acquisition of resistance to apoptosis and anoikis is also postulated to be a feature acquired during this stage, but still debatable as a feature of EMT [5,6,8].

According to the above mentioned authors, a typical EMT is when all these steps are fully achieved, whereas partial EMT is when at least one step is not fully or partially achieved. In this thesis, since we will only focus on the cell adhesion

changes during EMT, we consider EMT-like events that can result in phenotypes with adhesion features characteristic of Mesenchymal cell or partial EMT phenotypes.



**Figure 2. EMT archetype and Epithelial plasticity.** During EMT, epithelial cells lose their apico-basal polarity. Tight junctions which typically maintain apicobasal polarity dissolve allowing the mixing of apical and basolateral membrane proteins. Adherens and gap junctions are disassembled and cell surface proteins such as E-cadherin and epithelial-specific integrins (green) are replaced by N-cadherin and integrins specific to extracellular components (blue). The actin cytoskeleton is remodelled into stress fibres which accumulate in areas of cell protrusions. The epithelial intermediate filaments, cytokeratins, are replaced by vimentin. Meanwhile, the underlying basement membrane is degraded and the cell invades and moves into the surrounding stroma, devoid of cell-cell contacts. Figure (adapted) and legend reproduced from [5,6].

### 2.3. Partial EMT phenotypes

Phenotypes that conserve cell-cell adhesion during EMT (incomplete step 2 of EMT) are often the result of incomplete genetic program (partial EMT) [5,8,13]. Some of these phenotypes also acquire migratory capacity as Mesenchymal cells

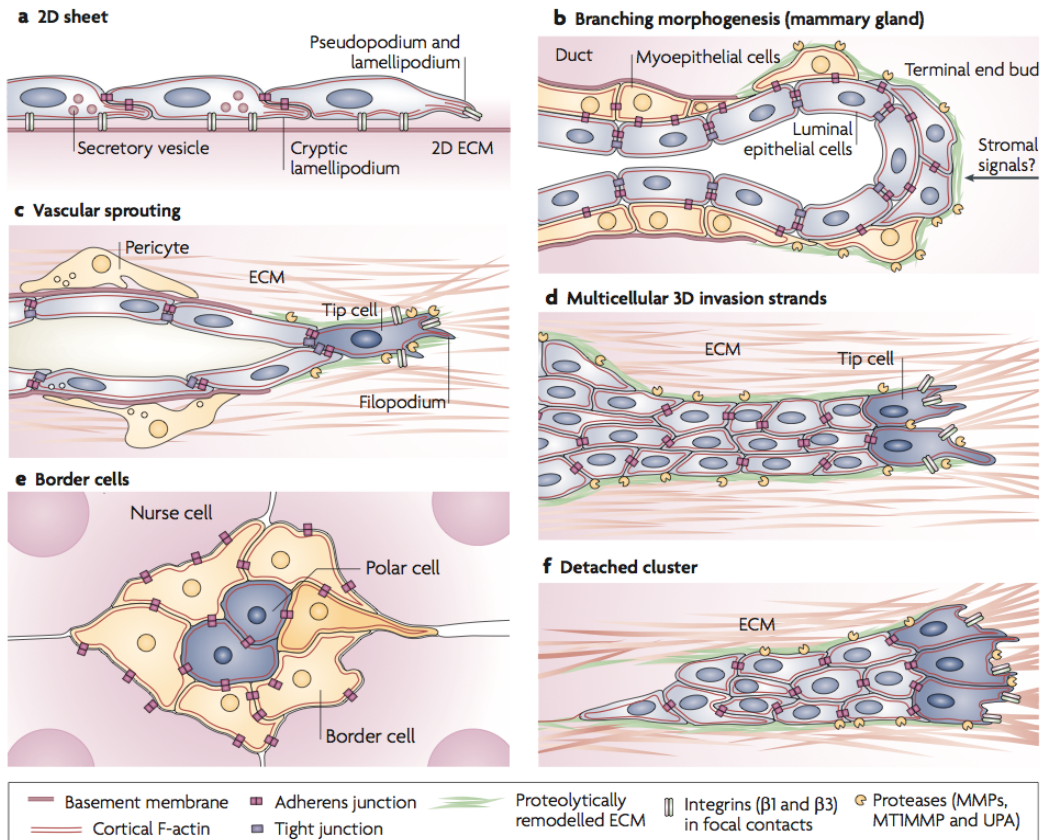


(complete step 3 of EMT, see section 2.2. for details), usually called as “Hybrid Epithelial-Mesenchymal phenotypes” [19,20]. Other “intermediate phenotypes” which do not acquire Mesenchymal migratory capacity can also result from partial EMT [5]. In this case, step 2 and 3 of EMT are incomplete (section 2.2.), where lamellipodia, filopodia, and invadopodia are absent or reduced in their capacity. In this thesis, we adopted the names “Hybrid” and “Intermediate” phenotypes for these two possible outcomes of partial EMT [5,20,21]. Together, Hybrid and Intermediate phenotypes can be part of a diversity of collective migratory behaviours (Figure 3) [19,20]. The observation of these phenotypes came from several *in vivo* and *in vitro* models which have been subject to EMT driving signals, suggesting that these phenotypes could result from partial EMT [5,19,22]. Collective migration phenotypes such as Hybrid phenotypes have been observed in 2D sheet movement of cells on an epithelial layer (Figure 3a). This type of collective behaviour was found in the intestinal epithelium and in epidermal keratinocytes during wound closure, suggesting it plays a role in tissue repair [5,22]. These cells are characterized by having cell-cell adhesion without tight junctions (incomplete step 1 of EMT), together with the acquisition of Mesenchymal-like membrane protrusions with lamellipodia and pseudopodia (step 3 of EMT) [5,22]. Hybrid and Intermediate phenotypes have been observed together in multicellular 3D clusters moving through a tissue (Figure 3b-f). These multicellular 3D clusters are composed by cells with distinct roles that lead to the classification into “leaders” (the cells located at the tip of the moving front) and “followers” (the cells located at the rear of the moving front) [22,23]. “Leaders” are characterized by having fully Mesenchymal migratory capacity with long membrane protrusions, driving the collective movement (step 3 of EMT), the Hybrid phenotype. “Followers”, on the other hand, are devoid of Mesenchymal migratory features such as the Intermediate phenotypes [22]. “Leaders” and

“followers” have been found to be organized into strands of cells or in small detached clusters (typically 6-8 cells) [22,24]. Collective migration of multicellular 3D clusters with both “leaders” (Hybrid phenotype) and “followers” (Intermediate phenotype) have been found in the morphogenic duct and gland formation during branching morphogenesis (Figure 3b) and in vascular sprouting during angiogenesis (Figure 3c) [5,22]. On these biological processes, the multicellular 3D clusters are characterized by strands of cells that maintain tight junctions and cell polarity, forming a tubular structure with an inner lumen [5,22].

Another example of 3D multicellular clusters is the migration of border cells observed in *Drosophila melanogaster* egg chamber during oogenesis (Figure 3e). This type of migration is characterized by an isolated small group of cells that migrate through the tissue without the typical migratory features of Mesenchymal cells (Intermediate phenotypes). In addition, these type of cells does not have tight junctions (step 1 of EMT), which ensures the characteristic loss of cell polarity and flexibility necessary for border cells migration. Interestingly, 3D solid strands (Figure 3d) and small clusters detached from these strands (Figure 3f) have been found in several types of carcinomas and implicated in invasion and metastasis formation [13,22,24]. These solid strands and small clusters move as poorly organized masses of carcinomas cells without the typical Epithelial cells polarity (dissolution of tight junctions, step 1 of EMT), features not observed in normal 3D collective migration such as vascular sprouting or branching morphogenesis [22,24]. Both solid strands and the small detached clusters are characterized by having cells typically differentiated into “leaders” and “followers” that resemble the Hybrid and Intermediate phenotypes, respectively. In carcinomas, Hybrid and Intermediate phenotypes are characterized by completed step 1 of EMT (Intermediate and Hybrid phenotypes) and step 3 in the case of the leaders (Hybrid phenotypes only) [6,24]. These observations lead to the idea

that EMT steps 2 and 3 are determinant for the stability of partial EMT phenotypes, especially in carcinomas.



**Figure 3. Types and variants of collective cell migration.** (a) Epidermal monolayer moving across a two-dimensional ECM substrate. Actin-rich pseudopodia and lamellipodia lead the migration with cells connected through adherens junctions. (b) Terminal end bud sprouting in the developing mammary gland during branching morphogenesis. Induced by stromal signals, the end bud extends from a duct through the protrusive movement of tight junction-connected luminal epithelial cells and loosely connected myoepithelial cells. (c) Vascular sprouting in newly forming or regenerating vessels. A tip cell with filopodial protrusions leads the migration through ECM. (d) Invasion of poorly differentiated multicellular masses and elongated strands in cancer. (e) Border cell cluster consisting of mobile outer cells and two less mobile polar cells migrating along cell-cell junctions of nurse cells in the *Drosophila melanogaster* egg chamber. (f) Collective invasion of detached cancer cells that are moving as a small cluster. F-actin, filamentous actin; MT1MMP, membrane type 1 matrix metalloproteinase (also known as MMP14); UPA, urokinase-type plasminogen activator. Reproduced figure and adapted legend from [22].

## 2.4. EMT in carcinomas

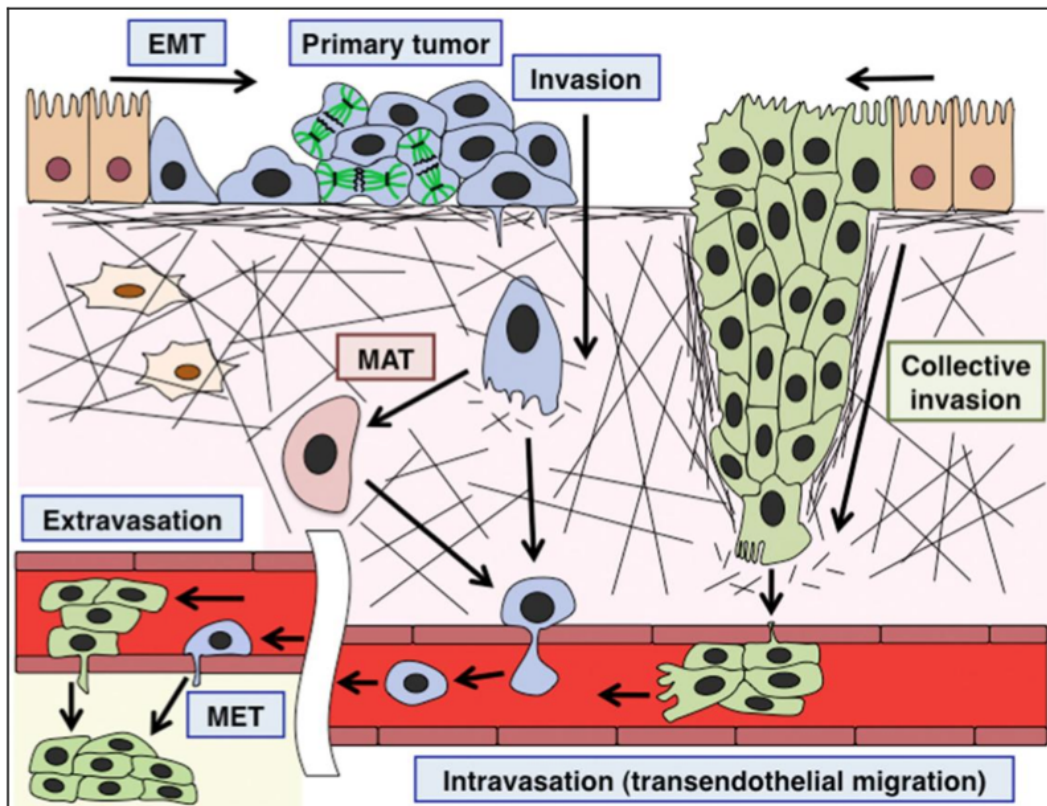
According to the National Cancer Institute (<https://www.cancer.gov>), carcinomas are a type of cancer that originates from a mutated Epithelial cell in the epithelial layer surrounding an organ, a gland or a body structure. Carcinomas are the most frequent type of cancers, affecting more than 10 million people in the world regardless of their sex or race [11,25,26]. Statistical studies in human populations also indicate that carcinomas such as the ones that develop in lung, breast, colon and stomach are the most deadly types of cancers, leading to around 1-2 million of deaths every year in Europe and US [25,26]. The poor survival prognosis of carcinomas has been associated with the acquisition of the capacity to invade and spread throughout other organs, designated by metastasis [11,27]. Invasion and metastasis are further considered as a hallmark of cancer progression which determines the malignancy of carcinomas [28]. Cancer therapy has very poor success when carcinomas start to spread, forming metastasis. This has motivated further cancer research in order to understand how metastasis is achieved and to find new therapeutic strategies [15]. Cancer researchers believe that during cancer progression, carcinomas acquire invasive properties by hijacking the EMT genetic program [5,7,13,15]. Although this is still debatable based on lacking clinical proof, several *in vitro* studies from mouse models and cell cultures have supported the role of EMT in invasion [13,15,29]. Researchers, further argue that only through undergoing EMT-like programs, carcinomas are able to metastasise [6,7,13]. Thus, understanding how to control EMT in the case of carcinomas could provide the means to fight metastasis, ultimately through the identification of new drug targets [15].

#### **2.4.1. Modes of invasion and metastasis**

As possible outcomes of EMT in carcinomas, Mesenchymal and Hybrid phenotypes have been associated with metastasis acquisition by ensuring single and collective migration capacity during invasion, respectively [24,30,31]. Amoeboid-like migration behaviour was also observed in a few breast cancers and small lung carcinomas, where cells migrate as isolated round shaped cells, without the formation of long membrane protrusions, often with chaotic movement and blebbing morphology (Amoeboid-like phenotypes) [30–32]. Metastasis of carcinomas is achieved through a complex and dynamical process of invasion, migration and interconversion between Epithelial, Hybrid, Mesenchymal and Amoeboid-like phenotypes [7,20,21,31,33]. In addition, the process involves multiple steps from the initial tumour site until the colonization of a secondary site (Figure 4). EMT provides one way to promote invasion, where the transformed cells either by Mesenchymal cells (complete EMT) or Hybrid phenotypes (partial EMT) degrade the basal membrane and escape the primary tumour site [24,30,31]. As cells migrate through the ECM, the Amoeboid-like phenotype can also arise from the process of Mesenchymal-to-Amoeboid transition (MAT) [31,33–35]. This process is controlled by GTPases Rac1 and RhoA, resulting in a rapid interconversion between Mesenchymal and Amoeboid-like phenotypes, ensuring an advantage in the migration through soft tissues as “pathfinders” [34,35].

To further advance in the process of invasion, carcinoma cells must gain the capacity to penetrate blood vessels through a process called intravasation [7,36]. Mesenchymal and Hybrid phenotypes are known to be more effective in the intravasation through the formation of actin-rich long protrusive invadopodia [37]. Then, carcinoma cells may enter the circulation and travel through the bloodstream as small multicellular 3D clusters (Hybrid) or as single cells

(Mesenchymal) [7,36]. *In vivo* studies have confirmed the presence of circulating cancer cells mainly as multicellular 3D clusters, suggesting that Hybrid phenotype is the main form of migration of carcinomas in the bloodstream [24,30,38]. To invade secondary sites and initiate the colonization process, the circulating carcinoma cells must exit the blood vessels through a process called extravasation [7,36]. Next, carcinoma cells penetrate and invade the basal membrane of the new organ until reaching the epithelial layer [7,36].



**Figure 4. Modes of invasion of carcinomas and metastasis acquisition.** Complete EMT mediated single cell invasion (left) and incomplete EMT mediated collective invasion (right). Normal epithelial cells (orange cells). Mesenchymal cells that result from complete EMT of a primary tumour (blue cells). Collective cell migration and invasion resulting from incomplete EMT with Hybrid and Intermediate phenotypes (green cells). Figure reproduced from [31].



Finally, to integrate and colonize the epithelial layer of the new organ, the invading Mesenchymal cells are thought to undergo MET [7,36]. For the colonization of multicellular 3D clusters in the epithelial layer, it is believed that they undergo the process of Hybrid-to-Epithelial transition (HET) [20,21,33]. Cancer researchers postulate that these two alternative modes of migration dynamically cooperate to achieve a more effective tumour invasion and colonization [20,38]. Moreover, this dynamics is also believed to be involved in the resistance to drugs and recurrence of carcinoma *in situ* upon removal. Therefore, the understanding of how to control the formation of these phenotypes and their interconversion in the metastasis context it is still an important challenge in the cancer field [6–8].

#### **2.4.2. Cell adhesion role in carcinoma invasion**

Integrin mediated cell-cell adhesion and E-cadherin mediated cell-cell adhesion are two properties that change during EMT and play a critical role in carcinomas invasion (Figure 5) [16,39,40]. During cancer invasion, integrin mediated cell-matrix adhesion is characterized by fast cycles of assembling and disassembly of focal adhesions, which lead to contraction of the membrane protrusions, generating the front-rear pulling force necessary for producing movement along ECM [31,40–42]. These focal adhesion dynamics have been considered as characteristic features of Mesenchymal and Hybrid phenotypes [31,39,40]. The assembling of focal adhesions involves clustering of integrins at a given membrane location (Figure 5), where more than 800 adapter and signalling proteins are recruited (adhesome) [43]. These structures ensure a strong integrin mediated adherence to specific ECM components such as laminin, collagen I and fibronectin, necessary for the pulling force [43–45]. On the adhesome, specifically through Talin binding, actin is polymerized into F-actin stress fibres by small Rho

GTPases [39,42,45]. Then, F-actin causes the elongation of membrane protrusions with the necessary rigidity for invading the basement membrane [39,42,45]. In migrating cells, the high rates of formation and turnover of focal adhesions (high focal adhesion dynamics) potentially can result in faster movement of cells, a property often correlated with cancer invasion [30,40,41]. The dynamics of focal adhesion formation and turnover have been found to depend on the activation of integrin signalling by ECM signals and growth factors [18,46–48]. The activation of FAK-SRC complex and PAK downstream of integrin signalling have been associated with both high turnovers of focal adhesions and cancer invasion [49–53]. Moreover, this has been also associated with the over-expression and over-activation of  $\alpha 5 \beta 1$  integrins on focal adhesions, ensuring high contractile forces during cancer migration [47,54].

In contrast, typical Epithelial cells lack high focal adhesion dynamics and are characterized by permanent integrins mediated attachment to the ECM (strong cell-ECM adhesion), which plays other functional roles such as Epithelium fixation and preventing apoptosis (anoikis) [5,45,55]. For phenotypes with amoeboid-like migration, the cell-ECM adhesion is absent or very reduced, allowing them to move freely on soft tissues [31,34].

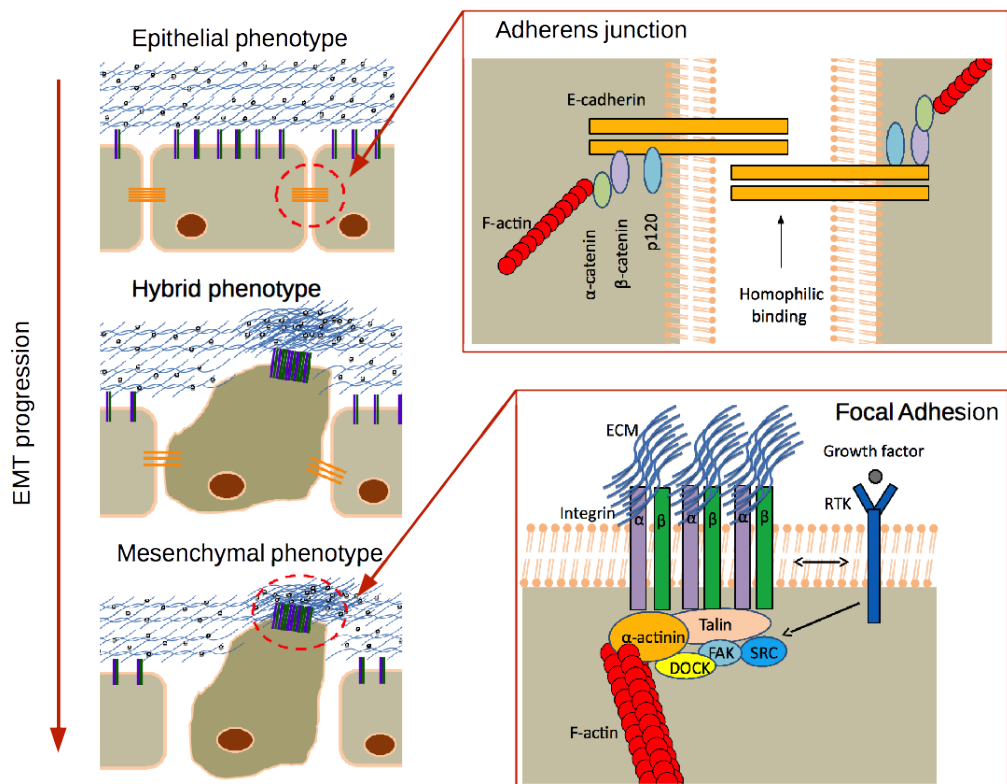
On the other hand, cell-cell adhesion through E-cadherin was found to be critical for the maintenance of the adherence between cells during collective migration of cancer cells [16,39]. This places E-cadherin mediated cell-cell adhesion at the frontier between single and collective forms of invasion [24,31,39]. Thus, this type of adhesion is a common feature of both Epithelial and Hybrid phenotypes, and absent in Mesenchymal and Amoeboid-like phenotypes [16,31,39]. E-cadherin is a transmembrane protein located at the adherens junction and holds cell-cell adhesion through homophilic binding with E-cadherin molecules of the neighbouring cells (Figure 5) [56–59]. E-cadherin binds to catenins  $\alpha$ ,  $\beta$  and  $\delta$ ,



forming a complex that participates in the maturation of cell-cell adhesion [16,57,58]. First, E-cadherin binds with catenin  $\delta$ , also called p120, which was demonstrated in several Epithelial cell lines to stabilize E-cadherin at the surface, increasing the abundance of E-cadherin able to participate in the cell-cell adhesion [60–62]. Secondly, the maturation of cell-cell adhesion proceeds after the binding of E-cadherin with  $\beta$ -catenin [63,64]. Once the E-cadherin- $\beta$ -catenin complex is formed, the docking for  $\alpha$ -catenin is possible, allowing G-actin to be polymerized into F-actin by small Rho GTPases, linked to the E-cadherin- $\beta$ -catenin- $\alpha$ -catenin complex [57,63,65]. Next, the F-actin cables re-enforce the strength of cell-cell adhesion through the concentration of adhesion molecules in the lateral side of cell-cell contacts, in particular, E-cadherin and Nectins [57,63,65,66]. For these reasons, the E-cadherin- $\beta$ -catenin-p120 complex is considered as a proxy for the cell-cell adhesion strength [16,65]. The strength of adhesion can be compromised through phosphorylation of p120 and  $\beta$ -catenin by tyrosine kinases activated during EMT such as SRC and MET, demonstrated *in vitro* to be capable of dissolving the E-cadherin-catenins complex [67–69].

On the other hand, specific receptor-type protein tyrosine phosphatases (RPTPs) are capable to dephosphorylate catenins which can recover the E-cadherin binding forms of p120 and  $\beta$ -catenin [70,71]. This suggests possible intermediate degrees of cell-cell adhesion for Intermediate and Hybrid phenotypes by modulation of RPTP and tyrosine kinases activity. Indeed, Epithelial cells can exhibit stronger cell-cell adhesion in comparison to collective migrating 3D cell clusters during invasion [24,31,39]. This is in part explained by a partial formation of the E-cadherin-catenins complex resulting from phosphorylation of p120 or  $\beta$ -catenin [16,57]. For example, in *Drosophila*, border cell migration occurs without  $\beta$ -catenin at adherens junctions [72]. Moreover, the loosening of cell-cell adhesion in collective migration during cancer invasion is thought to be a

necessary trait to allow intravasation and extravasation [31,39]. Taken all together, the combination of E-cadherin mediated cell-cell adhesion and focal adhesion dynamics are two properties able to distinguish invasive from non-invasive cancer cells and their correspondent modes of invasion (summarized in Table 1) [31,39]. Moreover, these adhesion types are critical properties necessary for cancer invasion and suitable traits for the characterization of the Mesenchymal, Hybrid and Epithelial phenotypes [31,39].



**Figure 5. Critical cell adhesion changes during EMT progression.** Cell-cell adhesion by adherens junctions and cell-matrix adhesion through the formation of focal adhesion. Representation contains only the critical components of the adherens junction and focal adhesion, based on the figure in [39].

**Table 1.** Cell adhesion properties of invasive and non-invasive phenotypes.

Phenotypes	Focal adhesion dynamics	Adherens junctions	Invasiveness	Behaviour
Epithelial	Reduced	Fully mature	Not invasive	Collective
Intermediate	Reduced	Partial	Reduced	Collective
Amoeboid-like	Reduced	Reduced	Mild	Single
Hybrid	High	Partial	High	Collective
Mesenchymal	High	Absent	High	Single

### 2.4.3. The role of the microenvironment

The microenvironment is a complex and dynamic system including signals from diffusible molecules, biophysical properties in the ECM, and cell-cell contacts from neighbouring cells [30,73–75]. In this thesis, we consider neighbouring cells as the adjacent cells to a cell in the Epithelial layer which is able to provide juxtacrine signals (cell-cell contact dependent) [73–75]. In the case of signals from the diffusible molecules in the microenvironment, we considered that they can come from endocrine, paracrine or even autocrine origin. Currently, the tumour microenvironment is postulated to play a fundamental role in driving metastasis of carcinomas [24,30,75,76]. Although mutation during cancer progression is thought to contribute to the acquisition of the metastable phenotype, the changes in the tumour microenvironment are believed to be required to trigger the transition between phenotypes with and without migration capacity (Mesenchymal/Hybrid and Epithelial, respectively) [20,30,75]. Through the last 2 decades, multiple signals from the microenvironment have been considered as hypotheses to drive the migration of carcinomas [74–76]. Among them, the addition of growth factors such as TGF- $\beta$ , EGF and HGF in the microenvironment (diffusible molecules) has been reported to be able to trigger EMT on carcinoma cell lines in cell culture conditions [77–79]. *In vivo*, the up-regulation of these growth factors in the tumour microenvironment can be

achieved through the secretion by the tumour cells or by recruited fibroblasts and macrophages during chronic inflammation [80–83]. These 3 cell types can also secrete to the tumour microenvironment collagen type I and fibronectin, remodelling the ECM composition [18,81,84]. These changes in ECM composition cause an increase in the ECM stiffness (biophysical property), which have been also demonstrated to cause the down-regulation of E-cadherin (EMT) in ovarian and squamous carcinoma cell culture conditions [85–87]. Due to accelerated tumour growth, oxygen consumption can lead to the state of hypoxia in the tumour microenvironment [76,84]. Hypoxia has been found to further promote the metastasis outcome in cooperation with the ECM composition, suggesting that multiple signals can synergistically lead to the metastasis outcome [84]. Chronic inflammation has been also associated with hypoxia and tumour invasion by leading to the collagen type I induced ECM stiffness and by the production of ROS, TGF- $\beta$ , HGF and cytokines [81,84,88,89]. Among cytokines, IL6 is hypothesised to play a role in cancer invasion based on reported evidence for its capacity to stimulate EMT in breast cancers cell lines, and because it is frequently overexpressed in the tumour microenvironment [90]. The endogenous production of ROS, not only drives mutation acquisition but also triggers HIF1 expression, which has been involved in EMT stabilization [81,91–93]. Recently, signals from the neighbouring cells mediated by Delta-like ligands activation of Notch receptors have been considered as juxtacrine signals in the tumour microenvironment that contribute to the invasiveness of carcinomas [20,94]. Unfortunately, the microenvironment signals that drive MET and stabilize Hybrid phenotypes (incomplete EMT) are still unknown [74,75]. Therefore, the microenvironment changes that lead to the switching between these phenotypes are yet to be characterized [20,30]. Thus, how the microenvironment controls the

switching between Epithelial, Hybrid and Mesenchymal phenotypes is still an open question in the cancer field.

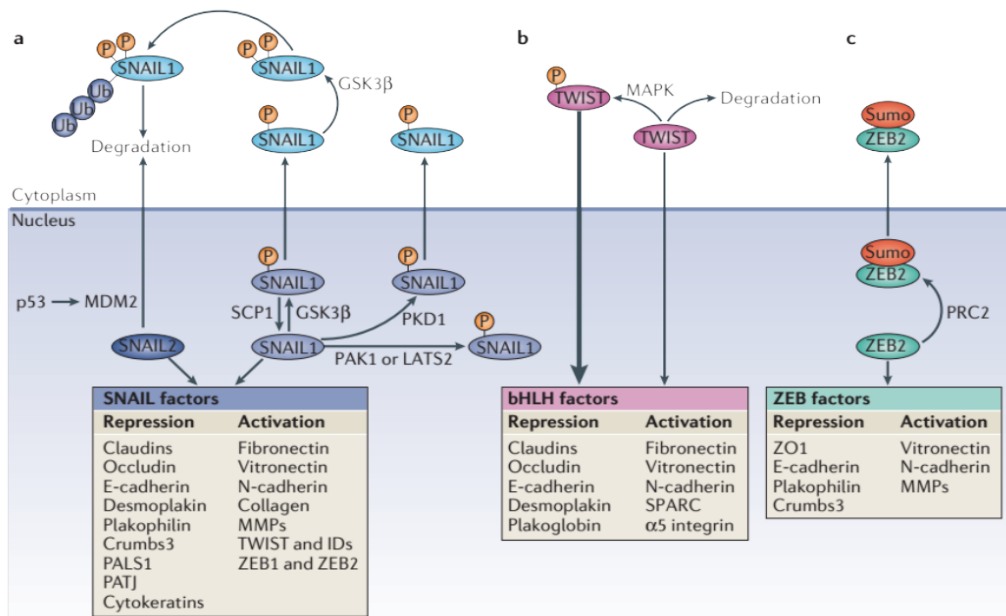
Interestingly, MET was observed to be induced *in vitro* with co-cultured normal Epithelial cell lines with Mesenchymal-like carcinoma cell lines, suggesting that cell-cell contact signals by neighbouring cells could trigger MET [95]. A vast number of cell-cell contact molecules have been found to trigger signalling between cells, opening new possibilities in the cancer field to understand how these signals could modulate cancer invasion dynamics [73,96]. In this thesis, we will only focus on 3 cell-cell contact signals ligands, the Notch, Receptor Protein Tyrosine Phosphatases (RPTP) and the membrane receptor FAT4 ligands. These last 2 cell-cell contact signals are quite interesting because are reported to inhibit particular EMT signalling pathways, suggesting that they potentially could prevent EMT [70,97]. In addition, these two receptors are good candidates because are ubiquitous expressed in most cell types and are reported to be strongly activated only by cell-cell contacts in high cell confluence conditions, i.e. high numbers of neighbouring cells (high cell density conditions) [98–101]. However, the effects of these signals are quite difficult to predict since multiple signals are present in the tumour microenvironment making this depend on the regulatory mechanisms involved. Based on these reasons, we chose to focus on the analysis of RPTPs and FAT4 as two possible hypotheses that could explain MET.

## **2.5. Transcriptional regulation of EMT**

The transcription of genes during EMT has been found to be under a complex and intertwined regulation, involving multiple transcriptional, post-transcriptional and even alternative splicing mechanisms [10,102]. During the last 2 decades, more than 21 transcriptional factors have been identified to be up-regulated in transformed cell lines and involved in the regulation of critical genes during EMT

[8,10,102]. Among them, Snail, Zeb and Twist family of transcription factors have been found to be the master regulators of EMT, which simplified the view of EMT regulation [10,102,103]. In particular, Snail family member Snail 1, because it controls the expression of both Zeb and Twist families [103–105]. These key transcription factors have been found to directly down-regulate several genes of key molecules associated with Epithelial features, and up-regulate several genes of key molecules associated with Mesenchymal features [10,102] (Figure 6). One of the most relevant targets of Snail, Zeb and Twist family is the E-cadherin gene (CDH1), considered as the hallmark of EMT [10,102,103,106].

Studies in cancer cell lines have demonstrated that Snail, Zeb and Twist directly bind to the DNA sequences of the E-box of E-cadherin gene promoter (CDH1) repressing the gene expression [10,103,107]. The activity of Snail, Zeb and Twist has been demonstrated to be under a complex post-translational regulation that results from upstream signalling, involving multiple phosphorylations, lysine oxidation and ubiquitination (Figure 6) [10,102]. The transcriptional regulation of Snail and Zeb is also subject to epigenetic control by histone modifications through several deacetylases, demethylase and methyltransferases, further increasing the layer of complexity [10,102]. As negative regulators of EMT, several small non-coding RNAs (micro RNAs) have been identified to play a tumour suppressing role by selectively bind to mRNA of Snail, Zeb and Twist, targeting them to degradation [10,102]. The miR200 family is the most studied and important regulatory micro RNAs based on its capacity to counteract invasion of several cancer models and recover of Epithelial phenotype [108–111]. However, the expression of Snail and Zeb family is also able to repress the expression of these micro RNAs resulting in feedback which can drive different cell fates and explain bi-stability of Epithelial and Mesenchymal phenotypes [10,102,112,113].



**Figure 6 Roles and regulation of major EMT transcription factors.** EMT is driven by SNAIL, zinc-finger E-box-binding (ZEB) and basic helix-loop-helix (bHLH) transcription factors that repress epithelial marker genes and activate genes associated with the mesenchymal phenotype. Post-translational modifications regulate their activities, subcellular localization and stability. **(a)** Glycogen synthase kinase-3 $\beta$  (GSK3 $\beta$ ) phosphorylates (P) SNAIL1 at two motifs; phosphorylation of the first motif facilitates the nuclear export of SNAIL1, and phosphorylation of the second motif enables the ubiquitin (Ub)-mediated degradation of SNAIL1. Phosphorylation of SNAIL1 by protein kinase D1 (PKD1) also leads to its nuclear export. Conversely, phosphorylation of SNAIL1 by p21 activated kinase 1 (PAK1) or large tumour suppressor 2 (LATS2), or dephosphorylation of SNAIL1 by small C-terminal domain phosphatase1 (SCP1) promotes the nuclear retention of SNAIL1 and enhances its activity. SNAIL2 is degraded as a result of its p53-mediated recruitment to the p53-mouse double minute 2 (MDM2) complex. **(b)** TWIST is phosphorylated by the MAPK p38, JUN N-terminal kinase (JNK) and ERK, which protects it from degradation, and thus promotes its nuclear import and functions. **(c)** ZEB2 is sumoylated (Sumo) by Polycomb repressive complex 2 (PRC2) and subsequently exported from the nucleus, which reduces its activity as a transcription factor. E-cadherin, epithelial cadherin; ID, inhibitor of differentiation; MMP, matrix metalloproteinase; N-cadherin, neural cadherin; PALS1, protein associated with Lin-7 1; PATJ, PALS1-associated tight-junction protein; SPARC, secreted protein acidic and rich in Cys; ZO1, zonula occludens 1. Figure and legend reproduced from [10].

## 2.6. Signalling pathways involved in EMT

Multiple signalling pathways have been found to cooperate and participate in the regulation of EMT [10]. Among them, the transforming growth factor- $\beta$  (TGF- $\beta$ ) signalling has been considered as the major driver of the initiation and progression of EMT [10,102,114,115]. Mechanistically, TGF- $\beta$  signalling drives EMT through the crosstalk with multiple signalling cascades encompassing

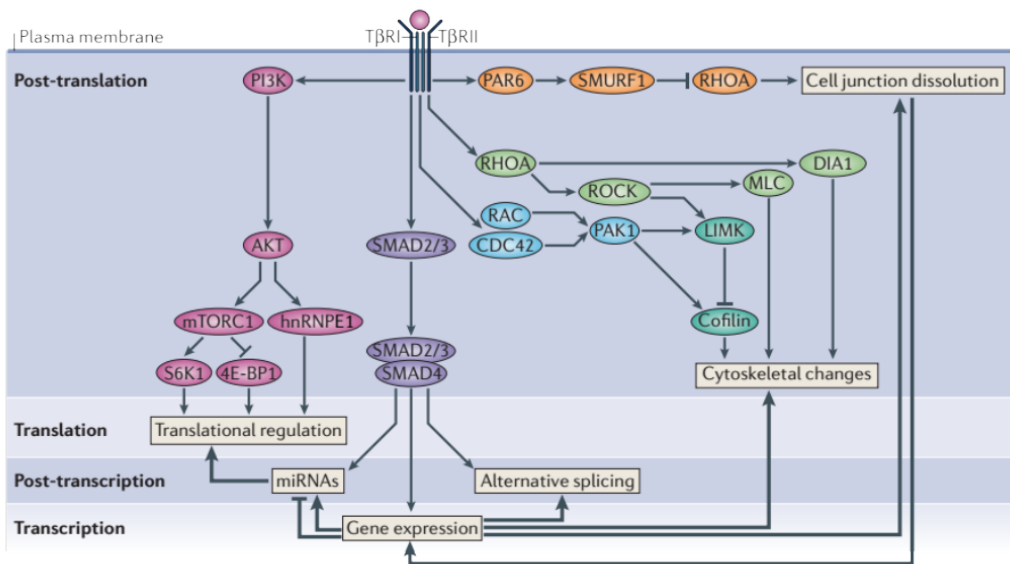
several transcriptional, post-transcriptional, translational and even post-translational mechanisms (Figure 7). TGF- $\beta$  signalling is initiated by the binding of TGF- $\beta$  protein family to the TGF- $\beta$  membrane receptors type I and type II, which in turn activates the SMAD dependent and independent signalling cascades, respectively [10,115,116]. Activation of TGF- $\beta$  receptors induce the first step of EMT through the activation of RHO-GTPases signalling, which causes subsequent cytoskeleton changes, dissolution of tight junctions and finally results in the loss of cell polarity [6,10]. Canonically, TGF- $\beta$  induces the activation of SMAD2, SMAD3, and SMAD4 into a complex that undergoes nuclear translocation and then participates in the transcriptional activation of several genes associated to the initiation of EMT such as the master EMT regulators Snail1 and Zeb [107,114,117]. Independently of SMAD signalling, TGF- $\beta$  can activate PI3K-AKT signalling resulting in the post-translational modifications necessary to trigger the translation or repression of additional genes associated with EMT progression [9,10,118]. For example, PI3K-AKT signalling is reported to stabilize Snail1 expression and therefore the Mesenchymal phenotype [10,119,120]. This is through AKT mediated inhibition of GSK3 $\beta$  by a phosphorylation mechanism, preventing Snail1 proteolysis via Snail phosphorylation by GSK3 $\beta$  [119,120]. PI3K-AKT signalling pathway also participates in the activation of the anti-apoptotic genes necessary to gain resistance to anoikis during migration and enhancing the expression of Snail1 through the activation of NF $\kappa$ B [9,10,118,121].

Several other signalling pathways cooperate during EMT progression to complete the EMT program (Figure 8). One of these signalling pathways is the canonical MAPK (RAS-RAF-MEK-ERK), reported to be fundamental for both initiation and progression of EMT in Epithelial cell lines [10,122]. Activation of MAPK signalling was found to be involved in EMT by the stimulation both Snail1 and Snail2



expression in several Epithelial cell lines through the activation of Egr-1 and AP-1 transcription factors [79,122,123]. Multiple receptor tyrosine kinases (RTKs) have been reported to activate canonic MAPK, including the receptors of several growth factors such as TGF- $\beta$ R, EGFR and MET [10,79,83,107]. Activation of integrins constitutes another alternative way for the activation of MAPK signalling, leading to EMT activation [10,46,124]. Interestingly, integrin and growth factor signalling were found to crosstalk and enhance the migration capacity of transformed cancer cells [46–48,125]. In addition, integrin signalling through the activation of ILK has been found to crosstalk with PI3K-AKT signalling in several carcinoma cell lines, suggesting this is an alternative way to lead to EMT progression [10,126–128]. Other MAPK members such as JNK and p38 have been found to be activated by several integrins and growth factors receptors such as TGF- $\beta$ R, also capable of stimulating the EMT progression [10,124,129].

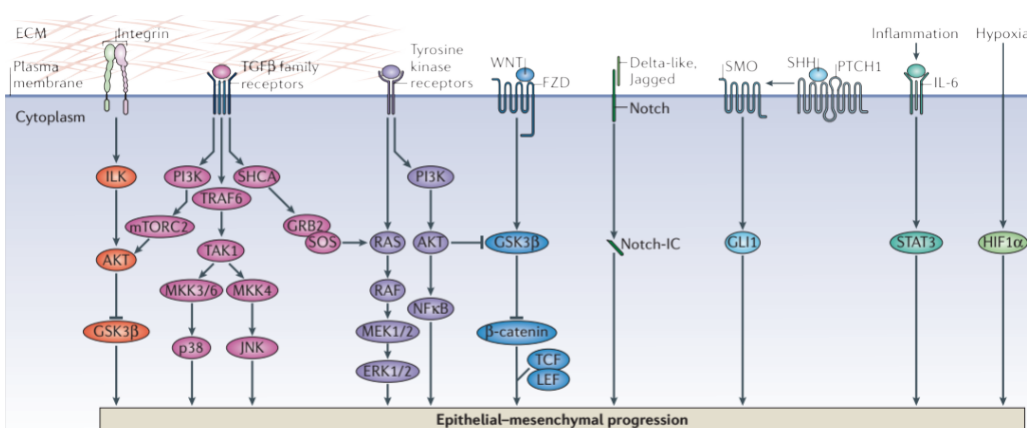
Activation of the canonic Wnt and Sonic hedgehog (SHH) developmental signalling pathways is another alternative way to drive EMT [10,130,131]. Wnt signalling mediates the nuclear translocation of  $\beta$ -catenin through the inhibition of GSK3 $\beta$  and  $\beta$ -catenin destruction complex [130,132]. Once  $\beta$ -catenin is in the nucleus, it binds to TCF/LEF super-family of transcription factors and activates the gene expression of Snail2 [107,133]. Recently, the joint activation of Wnt and SHH in hepatocyte cell lines demonstrated a positive feedback of Wnt signalling that sustained the autocrine activation of Wnt and TGF- $\beta$  signalling, further explaining the autocrine stabilization of the Mesenchymal phenotype in other cell lines [131,134]. The activation of TCF/LEF and SMAD gene expression in EMT progression have been also found to depend on the Hippo signalling through the activation of nuclear translocation of YAP and TAZ upon releasing LATS inhibition [135–137].



**Figure 7 Regulatory mechanisms of TGF- $\beta$  induced EMT.** TGF- $\beta$  complex type I and II receptors (T $\beta$ RI and T $\beta$ RII) activate SMAD2 and SMAD3, which then bind with SMAD4. The SMAD complex translocates to the nucleus and cooperates with transcription factors in the inhibition or activation of target genes decreases the expression of the epithelial splicing. TGF $\beta$  can also induce non-SMAD signalling pathways activating: PI3K–AKT–mammalian TOR complex 1 (mTORC1) signalling; RHO-GTPases through phosphorylation of PAR6–SMURF1 signalling leading to tight junctions dissolution; Diaphanous (DIA1) and also RHO-associated kinase (ROCK) to promote actin reorganization by the phosphorylation of myosin light chain (MLC) to activate LIM kinase (LIMK) and thus inhibit cofilin. RAC and CDC42 also participate in cytoskeletal changes through p21 activated kinase 1 (PAK1) and direct the formation of lamellipodia and filopodia. 4E-BP1, eukaryotic translation initiation factor 4E-binding protein 1; S6K1, ribosomal S6 kinase 1. Figure reproduced and legend adapted from [10].

Experiments with cancer cell lines have shown that the over-activation of Notch, STAT3 and HIF1 $\alpha$  signalling can further stimulate EMT leading to enhanced invasion capacity [10,89,92–94,138]. In addition, synergistic effects have also been found between Notch/HIF1 $\alpha$  and STAT3/HIF1 $\alpha$  in breast cancer and prostate cell lines suggesting cooperation between these signalling pathways for carcinoma invasion capacity [89,92,93]. In breast cancers, Notch signalling activation was found to directly activate Snail2, suggesting it can initiate EMT by

itself [139]. However, Notch was also found to crosstalk with TGF- $\beta$ , MAPK and AKT signalling by overexpressing both NF $\kappa$ B and EGFR, further increasing the complexity of the enhanced EMT stimulation in cancer [94,140,141]. Moreover, STAT3 and HIF1 $\alpha$  signalling are capable of enhancing EMT through the activation LIV1 and LOX expression respectively, two alternative ways of stabilizing Snail1 independently of GSK3 $\beta$  activity [92,93,142]. Taken all together, multiple and intertwined signalling pathways make the understanding of the regulation of EMT a complex and challenging scientific problem.



**Figure 8 Signalling pathways involved in EMT progression.** Integrin induced signalling (orange), SMAD independent TGF $\beta$  signalling (pink), Receptor tyrosine kinase induced signalling (purple), Notch signalling (dark green), Wnt signalling (dark blue); Sonic Hedgehog signalling (SHH in light blue); STAT3 signalling (STAT3 in green) Hypoxia inducing factor 1  $\alpha$  signalling (HIF1 $\alpha$  in light green), AKT signalling (PI3K-AKT), MAPK signalling (RAS-RAF-MEK1/2-ERK1/2). Additional abbreviations: ECM (extracellular matrix); FZD (frizzled); ILK (integrin-linked kinase); MKK (MAPK kinase) ; mTORC2 (mammalian TOR complex 2); Notch-IC (intracellular fragment of Notch);JNK; (c-JUN N-terminal kinase); SHCA (adaptor protein SRC homology 2 domain-containing-transforming A); GRB2 (growth factor receptor-bound protein 2); SOS (son of sevenless); TRAF6 (TNF receptor-associated factor 6); TAK1 (TGF $\beta$ -activated kinase 1); GSK3 $\beta$  (glycogen synthase kinase-3 $\beta$ ); TCF/LEF (transcription factors lymphoid enhancer-binding factor LEF and T cell factor TCF); PTCH1 (patched 1); SMO (smoothened); GLI1 (glioma 1); IL-6 (interleukin-6); JAK (Janus kinase); STAT3 (signal transducer and activator of transcription 3); NF- $\kappa$ B (nuclear factor- $\kappa$ B).. Figure reproduced from [10].

### 3. Addressing EMT through a Systems Approach

Systems biology is a multidisciplinary scientific field originated in the 60s to address biological-based complex problems with an integrative thinking (holistic),

instead of the traditional reductionist approach [143–146]. Systems biology encompasses a wide range of methodologies and concepts that came from computer science, mathematical modelling, enzyme kinetics, control theory, and cybernetics [144,145,147]. Since the year 2000, the field has been more dedicated to the analysis of large networks, motivated by the emergence of large quantitative and qualitative data sets, generated from the always evolving experimental techniques in genomics and proteomics [143,145,148]. This forced the improvement of computational and modelling methods to deal with such large multi-scale datasets [145,148,149]. Since then, several frameworks in systems biology have been evolving, combining both theoretical and experimental methods in a dynamic cycle to discard hypothesis and generate new testable predictions (Figure 9). Currently, the systems approach is considered to be a fundamental asset for the understanding of complex and multi-variable problems in biology such as signalling networks in cancer and other diseases [148,149].

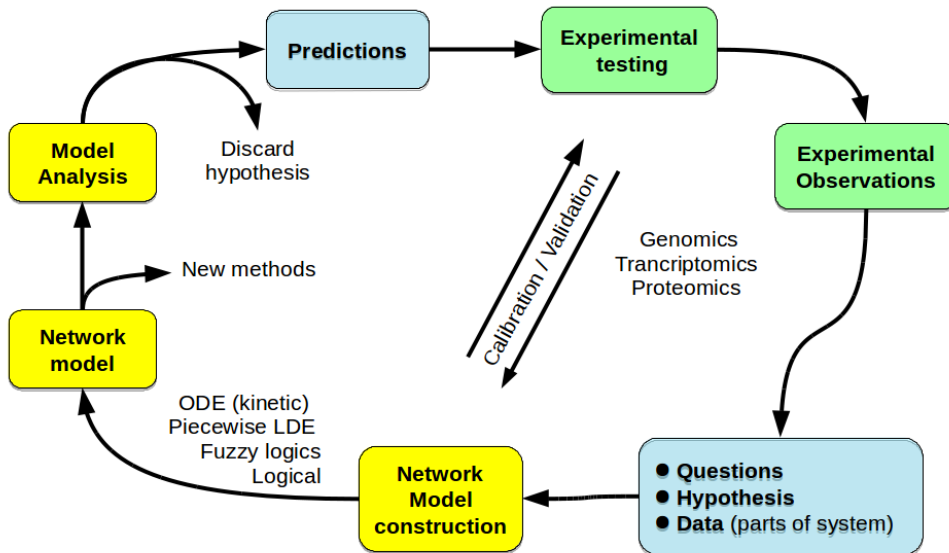
The regulation of EMT is indeed a multi-variable complex problem due to the involvement of multiple microenvironment signals, the intertwined nature of the involved signalling pathways, and the presence of multiple positive and negative feedback loops [8,10,102]. In addition, the complexity further increases in this system when mutational effects during cancer progression interfere with the regulation of EMT [28,150]. According to some cancer researchers, the integrative approach of systems biology could help to understand how EMT is controlled in the cancer context [8,20,151]. In particular, to identify drug targets that can prevent metastasis [149].

Experimental data concerning regulatory effects on EMT in normal and carcinoma cell lines have been accumulating in the literature during the last two decades at a rate superior to its integration and analysis [8,102]. Computational researchers have made efforts to store these data in publicly available databases

such as the KEGG PATHWAYS and the Atlas of Cancer Signalling Network (ACSN) [151,152]. Computational methods such as NaviCell and Cytoscape have been developed for the visualization, navigation and pathway analysis on these databases, which have already provided useful information [148,149,153]. However, the complexity and extent of the network of regulatory effects are humanly impossible to intuitively grasp behaviours from only looking at the descriptive network and analysing its patterns. Thus, the complexity of EMT regulation calls for mathematical modelling frameworks to integrate the documented regulatory mechanisms and analyse their dynamic behaviours.

Several continuous and discrete mathematical frameworks (Figure 9) have been used in systems biology to describe the regulation of biological processes and generate testable predictions [154,155]. The most popular are the kinetic and the logical modelling frameworks [155].

Briefly, the kinetic modelling framework relies on simulation or solving systems of ordinary differential equations (ODE) to generate continuous and quantitative changes of biological entities (variables) through time. In this framework, the rate laws ( $v$ ) of each variable ( $x$ ) is integrated in order of time ( $t$ ) with the form  $v = dx/dt$ , where  $\int dx dt = x(t)$ . For more details on kinetic modelling framework please see review in [155]. On the other hand, the logical modelling framework is based on a qualitative and discrete approach. In this approach, the changes of the variables ( $x$ ) through time ( $t$ ) are accounted through rules-based functions ( $F$ ) and an update scheme with the form  $x(t+1) = F(x(t))$ , where time has arbitrary units. A detailed description of logical modelling framework and comparison with kinetic modelling approach in the context of systems biology and cancer is presented in the next section and sub-sections.



**Figure 9. The systems biology cycle.** From questions in biology that include multiple hypotheses and large data sets, the systems approach gathers all pieces of information to construct a network model. The analysis of this network is through the association to mathematical modelling frameworks, which are selected to meet the questions and available data. The frameworks frequently used in systems biology are: Ordinary Differential Equations (ODE) based kinetic models; Logical models (Boolean and multi-valued); Fuzzy logic models; And piecewise Linear Differential Equations (LDE) models. To generate trustable predictions, models have been trained or calibrated with experimental data sets that result from genomic, transcriptomic or proteomic data. In some cases, models have been validated by comparing its results with independent experimental data. To analyse the models according to particular questions and dependent on the data itself, sometimes new analytical methods must be developed. Finally, the analysis of validated/calibrated network models provides a way to evaluate hypothesis and generate experimentally testable predictions, which can eventually be confirmed.

### 3.1. The logical formalism

The application of the logical formalism in biology started in the 60-70s with the work of Sugito, Kaufman and R. Thomas [156–158]. In modern systems biology, this framework is now being used with success in the analysis of large molecular regulatory networks (genes, proteins, microRNA's) to describe complex behaviours and cell fate decisions [155,159,160]. Briefly, the framework consists in the development and analysis of a logical model that describes one or more

phenomenological properties, assuming a non-linear regulation of its regulatory components, which depends only on the presence or absence of regulators [161]. A logical model is an abstraction of the regulatory network and is represented through a graph (regulatory graph). The graph edges reflect influence effects between the network components (nodes) that account for processes of activation or inhibition. In logical models, the nodes of the graph are the variables and are often binary (Boolean, ON/OFF), describing two qualitative states of biological activity or concentration for the respective molecular components. In alternative, the nodes can also be multi-valued, where multiple discrete and finite degrees of activity/concentration can be associated with molecular components. The behaviours of logical models are given by logical functions (also called logical rules) that define the evolution of the variables towards attractors, according to an update scheme. In general, there are two types of attractors in logical modelling, the point attractors and the cyclic attractors. Point attractors, are also called stable states (we will use more this term), which are fixed points where the functions are no longer updatable. This means that all variables evolved towards a stationary value, which is equivalent to the frequently called steady-state that result from kinetic models. In the case of cyclic attractors in logical modelling, the values of a particular sub-set of variables are always changing in  $t \rightarrow \infty$  with a defined oscillatory behaviour, whereas the remaining variables have fixed values (stationary). This behaviour is comparable with the limit cycle oscillatory behaviour observed in some kinetic models. Moreover, the logical framework introduces the notion of threshold values that define changes in the activity of the model components according to logical functions. Thus, logical modelling provides a discrete and qualitative analysis of molecular networks, rather than the commonly used quantitative and continuous ODE-based kinetic modelling [159,160].

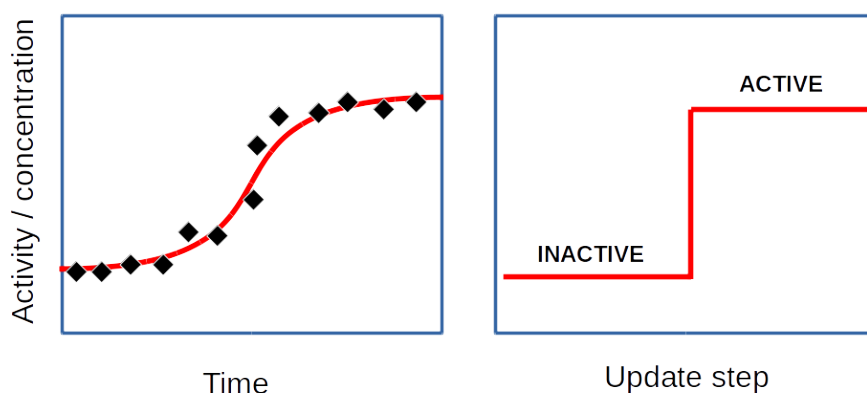
### **3.1.1. Advantages and limitations**

Logical modelling of molecular networks has advantages and limitations in comparison with commonly used continuous and quantitative modelling approaches (see [155] for a detailed comparison). Thus, the choice for the framework depends on the type of information available and the type of network to model. The most common approach in network modelling is through the development of kinetic models, which rely on detailed knowledge about the rate laws and kinetic constants of each individual process [155]. Logical modelling has the advantage of not requiring precise knowledge about the kinetics of the individual processes in the network [155,159,162]. This is particularly advantageous for most molecular networks containing signalling processes because their kinetic rate laws and kinetic parameters are usually unknown. Although kinetic parameters can be estimated through global optimization fitting methods to experimental data, it becomes very hard to find solutions in the case of large networks models [155]. In addition, the uncertainty around each estimated parameter scales up to the number of total parameters. Thus, logical modelling is then a suitable alternative for large regulatory networks such as signalling networks. Logical models only rely on qualitative information for the construction and testing of the resulting predictions. This type of data has the advantage of being easily obtained experimentally or by searching the literature. The reason is that most of the problems in biology were tackled with qualitative or semi-quantitative experimental techniques to report effects of inhibitions, gene knockout and overexpression over time. Thus, this type of methods are commonly available and a great number of reports with qualitative data are accumulated in the literature.

On the other hand, logical modelling has the limitation of being much less precise in describing individual biological processes in comparison with the kinetic



modelling approach (Figure 10) [155]. This is because logical modelling predictions are not continuous (discrete values) and therefore neglect the transitory behaviour between initial and final state of a process. Meanwhile, logical models can be as informative as kinetic models if we just want to know a global change (final - initial). Therefore, logical modelling are suitable to describe biological processes in threshold concentrations of a biological entity that result in a maximum response in a biological target such as biological processes that follow a sigmoid behaviour (Figure 10). For example, this is the case of gene expression (gene  $\rightarrow$  protein) and signalling processes above threshold concentrations of its activator/inhibitor (strong signal).



**Figure 10. Modelling the time-course kinetics using a kinetic model and the correspondent logical abstraction.** Red lines are theoretical values by both kinetic and logical models. Dots correspond to the experimentally observed values for an biological molecule in time. Most biological processes lead to a saturation of a product during its production or during a chemical transformation of molecule with time (sigmoid behaviour). This is when the rate of production of a biological chemical species equals the rate of consumption. For example: The production of a protein by its gene activation reaches its limit when its rate of production is equal to the rate of protein turnover; The phosphorylation of a protein by kinases (activation/inhibition) in signalling is limited by the constant protein concentration; Chemical reactions and molecular binding of proteins when they reach equilibrium.

Time is implicit in the logical framework through the variables update calls, which is defined by logical rules, without accessing the rates of the processes. This is a

good approximation if the rates of the processes are similar (in the same timescale). However, in particular conditions where the rates of the processes are different in orders of magnitude (different timescales), this may result in dynamics and regulatory feedback effects which are not biologically plausible. This limitation can be overcome by adding delays, probabilities or rules of priorities on logical models, given the availability of data that supports these features [155]. Finally, combinatorial explosion occurs as the number of states grows exponentially with the number of variables (model components). This hampers the analysis of large models (because the state space is too large). Moreover, model design may be hard due to the combinatorial explosion of the number of potential logical rules for components that have large number of regulators.

### 3.1.2. Model definition

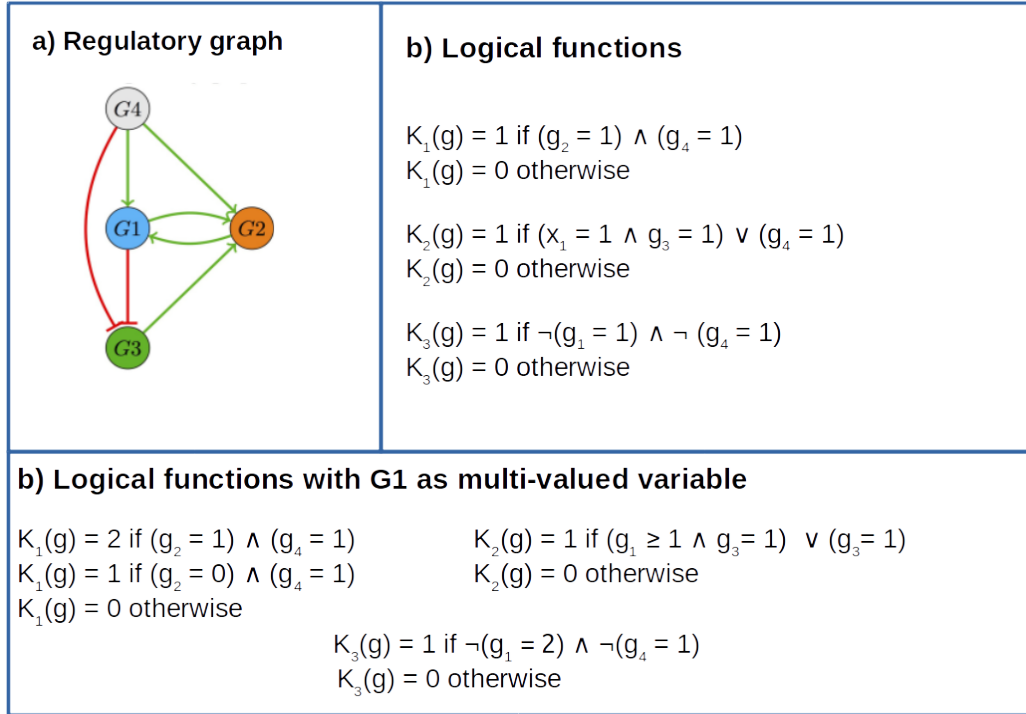
A Logical model is defined by a regulatory graph and associated logical functions (see the illustrative example in (Figure 11). According to the formal definition in [160], a logical model is a mathematical object defined by  $(\mathbf{G}, \mathbf{K})$ , where:

- $\mathbf{G} = \{g_1, g_2, \dots, g_n\}$  is the set of  $n$  network components, where each  $g_i$  ( $i = \{0, \dots, n\}$ ) is a variable that can adopt integer values from 0 to a defined maximum value ( $max_i$ ) as in the set  $\{0, 1, 2, \dots, max_i\}$ . Each value of  $g_i$  means a biological feature (e.g. concentration or activity degree). The vector  $\mathbf{g} = (g_1, g_2, \dots, g_n)$  contains all variables and characterizes a state of the model in the finite state space  $S$  given by the Cartesian product

$$\prod_{i=1}^n \{0, \dots, max_i\} .$$

- $\mathbf{K} = \{k_1, k_2, \dots, k_n\}$  is the transition function between model states, composed by a set of discrete logical functions  $k_i$ , each defining the value

of  $g_i$ , which depends on the current state  $g$  (and more precisely on the values of its regulators), which is depending on an initial state of  $g$ . Thus, the model behaviours are defined by the transition function  $k: S \rightarrow S$ , with each successor state of  $g$  given by  $K(g) = (k_1(g), k_2(g), \dots, k_n(g))$ .



**Figure 11 Formal logical model definition, an illustrative example.** (a) The regulatory graph that reflects the topology of the network, where the nodes are the variables of the model (network components). The edges represent the regulatory effects of activation (green) or inhibition (red). Gray node represents an input variable (no function attributed) and the remaining nodes are variables that are given by logical functions. (b, c) The logical functions are written for the regulatory graph variables to define the model dynamics, where **b** considers all variables Boolean and **c** the variable G1 as multi-valued. The mathematical symbols “ $\wedge$ ,” “ $\vee$ ,” and “ $\neg$ ” represent the logical operators AND, OR and NOT, respectively. Figure and legend adapted from [160].

The nodes of the graph correspond to the variables  $g_i$  in  $\mathbf{G}$  (network components) and the incoming edges the regulatory effects (activations and inhibitions). For

nodes with no incoming edges, the variables  $g_i$  are defined as model inputs and no logical functions are defined for these variables. In the case of nodes with no outgoing edges, these variables are usually considered as model outputs. Logical functions in  $\mathbf{K}$  are written accounting for the regulatory graph using conjunctive, disjunctive and negation between variables  $g_i$  (AND, OR, and NOT) [160]. Thus, for each variable  $g_i$  that has incoming edges (activations or inhibitions), a logical function is written considering the variables in  $\mathbf{G}$  that correspond to all incoming edges (regulators). Here, the rules are written by connecting the regulators through the logical operators (see Figure 11 for examples).

For multiple activations of a node, the AND rule is chosen when these regulators are required to be together, and the OR rule when each regulator can activate the target node independent of the status of the other regulators. For the inhibition of a node by a regulator, the rule is defined by using the NOT operator as the prefix. Therefore, multiple functions with distinct transition dynamics can be written for each graph, depending on the number of regulators. Thus, each graph can be associated with a family of logical models, and each  $\mathbf{K}$  corresponds to a unique regulatory graph [160]. Functions can be developed based on information from the literature, where the rules are chosen based on biological knowledge. Or in alternative, rules can be inferred from observed data through truth tables.

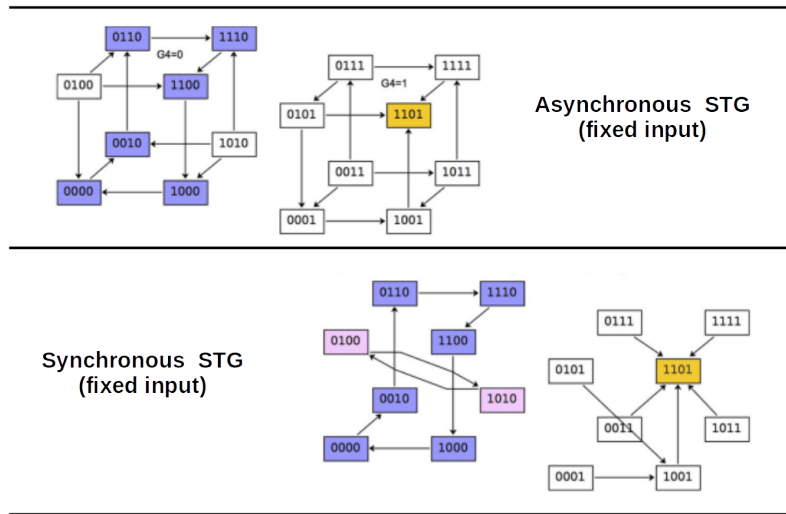
### 3.1.3. Model dynamics

A Logical model generates a discrete dynamics over  $S$ , which is defined by the transition function  $\mathbf{K}$  and depends on the initial state  $g$ . To generate a dynamics, if  $K(g) \neq g$  the function is called to update at least one variable in  $g$  to match the one(s) in  $K(g)$  [160]. Whenever  $K(g) = g$ , there is no call for update of variables, and this state is denominated a stable state (also called point attractor, which is a

stationary state). If not, variable updates generate transitions towards successive states. The dynamics is frequently represented by a State Transition Graph (STG), where the nodes of STG are the model states and each state is connected to its successor state through the edges [160]. The analysis of the STG provides the identification of the initial conditions and paths that lead to stable states or cyclic attractors (see examples in Figure 12). Biologically, stable states indicate a cell fate and cyclic attractors an oscillatory behaviour [159,160,162,163]. In the STG, the stable states, are identified by the nodes of the graph (states) that do not possess outgoing edges. For cyclic attractors, these are identified in the STG when the dynamics falls into a set of transient states that results in a closed loop of transitions (without outgoing edges). However, the STG of large network models can result in a huge graph that is impossible for a human to understand. As alternative, the Hierarchical Transition Graph (HTG) is a compact representation of an STG that hides most transitions, only showing the outgoing transitions that lead to stable states or cyclic attractors (for more details see [163] ).

The two most common forms to update variables are called the synchronous and the asynchronous update schemes [160]. These types of variables update can generate distinct dynamics (see examples in Figure 12). In the synchronous update, all variables in  $g$  change simultaneously in each update call. This update scheme originates deterministic dynamics, where one state can only generate a unique successor. In the asynchronous update, when more than one variable does not match its function ( $K(g) \neq g$ ), each node value is updated independently, generating non-deterministic concurrent states. Thus, asynchronous dynamics generates all possible concurrent transitions. Within the asynchronous updating scheme, probabilities can be used to randomly select the variable to update (Random asynchronous update), generating a particular dynamics and discarding

others [160]. This can be also applied in the case of networks, where the regulatory processes have distinct timescales such as in the case of phosphorylation of proteins and transcription of genes [160]. In these cases, defining high and lower probabilities can provide an approximation to the kinetics.



**Figure 12 Model dynamics, an illustrative example.** State Transition Graphs (STG) that represent the dynamics of the Boolean version of the model (Figure 11), starting from any state of the state space. On top, the dynamics with the asynchronous update and on the bottom with the synchronous update. Dynamics generated for each input condition fixed ( $G_4=0$  and  $G_4=1$ ). The values of the states  $g$  are denoted in each node of the STG as a string of the values with the following order:  $x_1, x_2, x_3$ , and  $x_4$ . Nodes in white are the transient states, yellow shows the stable states, blue shows a cyclic attractor and pink a new terminal cycle. Figure and legend adapted from [160].

In alternative, rules of priority can also be defined in the asynchronous update, where nodes are ranked into classes with an order of update [163,164]. Formally, for a set of variables in  $\mathbf{G}$  of size  $n$ , it is possible to define the priority classes  $C_1$ ;  $C_2$ ; ...,  $C_p$  with  $p \leq n$  [164]. Here, the variables that belong to the highest rank will be checked first for the update, followed by next rank only if no variables in the higher rank are call to update. This strategy is suitable to mimic molecular

entities that are transformed in different time scales and can be defined for the increase and decrease in activity of variables [164].

### 3.1.4. Methods and tools

Free software tools for the construction and analysis of logical network models have been developed by several groups [160,165]. Through the **Consortium for Logical Modelling and Tools (CoLoMoTo)**, a joint effort has been made to standardize and integrate most of the analytical methods in the already available software tools (Table 2) [165]. Extensive description and comparison of these methods are available in the CoLoMoTo web page (<http://colomoto.org/software>). Among them, the GINsim (Gene Interaction Network simulation) is a versatile software tool, exclusively dedicated to the analysis of regulatory networks using the logical framework (<http://ginsim.org>) [166]. With this tool, the regulatory graph can be design and the respective logical model can be formulated and analysed in a user friendly environment. Here, both Boolean and multi-valued variables are supported. The models can also be easily reduced in their number of components with a semi-automatized method described in [167]. The model dynamics can be generated under the synchronous, asynchronous update schemes, including the possibility of setting priority classes. GINsim is able to construct the STG with a reasonable size (in the order of a few million states), and also the respective HTG [163,166]. GINsim includes several methods to analyse model properties and supports multiple model formats enabling the usage in other software tools [163,166]. With GINsim, we can analyse static model properties such as; the analytical determination of all stable states regardless of their reachability conditions, and the identification of conditions for positive and negative functional circuits. To analyse the reachability of stable states, GINsim also recently included methods to estimate their probabilities such

as Avatar, MonteCarlo and FireFront. In GINsim, the user can define perturbations by simply fix the variables values and explore the changes in static and dynamic properties. For example, fixing a node to value 0 or 1 to mimic the effect of a putative knockout or over-activation of a gene, respectively. In GINsim, it is also possible to knockout or ectopically active the effect of specific interaction in a similar way. This is particularly interesting to study the importance of a regulatory interaction on the network properties.

**Table 2. Summary of logical modelling software tools.** Tools that belong to the CoLoMoTo [165]. Tools share the SBML format, which enables the integrative usage in the analysis of logical models.

Software tools	Description
<b>BoolNet</b>	R package for Boolean networks. Allows construction, analysis and simulation with several updating schemes.
<b>Cell Collective</b>	Web platform for Boolean networks. Allows construction, analysis and simulation.
<b>CellNetAnalyzer</b>	Toolbox (MATLAB) to provide a graphical user interface to explore structural and functional properties of metabolic and signalling networks.
<b>CellNOpt</b>	Open source application (R/BioConductor, Cytoscape, Python) for the training of logical network models with experimental data. Supports Boolean, Fuzzy logics and ODE formalisms.
<b>GINsim</b>	Java application for the analysis of Boolean and Multi-valued networks. Allows simulation of dynamics with several updating schemes.
<b>MaBoSS</b>	C++ developed application for the simulation of time Markov processes derived from Boolean logical models. Does not support SBML.
<b>BoolSim SQUAD</b>	Search and identify all type of attractors (stable and cyclic). Also allows the construction of continuous models from logical models.

### 3.1.5. Modelling cancer networks

Computational modelling of molecular regulatory networks using the logical formalism has been recently applied with success in cancer research [131,168–172]. Some modelling work has contributed to cancer research by finding new combinations of drug targets to fight cancer, proven to have a synergistic effect *in*



*vitro* or in a mouse model [170–172]. Other studies have contributed to cancer research by explaining particular behaviours or properties such as co-occurrence of mutational effects and invasive cell fate properties [131,168,169]. All these logical models have been constructed based on reports from literature, which supported each regulatory effect (interaction). Some authors developed their models for specific cancer types [131,168,170,172]. Others, have chosen to develop a generic model and applied it to a few cancer cell types [169,171]. These modelling choices depend on the biological question and also limited by the available knowledge of the biological system. Choosing a particular cancer cell type to focus simplifies model construction and validation if enough data is available for that cancer cell type. This choice allows exploring particular characteristics of a cancer type (“zoom in” approach), which may be useful to better understand a particular type of cancer and design specific therapeutic strategies. On the other hand, the choice of modelling generic properties (“zoom out” approach) is advantageous for searching conserved mechanisms of common cancer behaviours such as migration and metastasis acquisition. This may help to find critical targets for cancer control, which could improve the design of therapeutical strategies. However, the “zoom out” approach is more difficult in terms of model development and validation because of biological variation and the lack of comparable data for all cancer types.

To help the construction of networks, modellers have resorted to available signalling networks in databases such as KEGG PATHWAYS, ACSN and Reactome to identify relevant interactions, instead of inferring networks from experimental data. Logical functions have been derived by interpreting collected literature information instead of using fitting methods to experimental data. In alternative, multiple functions generating a family of models have been analysed to evaluate the robustness of generated predictions [171,172]. GINsim has been

extensively used for the construction of the regulatory network and for the generation of the model dynamics [168,169,171,172]. Other software tools like MaBoSS have been also resourced, in particular for the analysis of reachability properties [171]. Perturbations of the network nodes (components) have been fruitful to generate predictions and commonly resourced for the simulation of knockout, inhibitions and over-expression experiments used in model validation. Logical models in cancer have been validated by recapitulating phenomenological observations from a set of experiments on cancer cell lines [168–172]. In addition, logical models have been challenged to recapitulate some molecular changes in activity or concentration during such experiments. Used experiments consisted on the quantification of changes in phenomenological properties in a population of cells in culture with and without a given set of perturbations (e.g. gene knockout, gene over-expression and addition of growth factors), reflecting the experimental conditions. Most models have used experimental observations collected from literature articles or expression data from public databases instead of producing their own data. However, most modellers test their predictions *in vitro* using cell lines, which in some cases are in collaboration with experimental groups. For the comparison with logical modelling predictions, modellers often binarize the observed phenomenological changes to qualitative changes such as up-regulated (value 1) and down-regulated (value 0). To do so, modellers such as Cohen et al. resorted to specific software tools, others simply qualitatively interpret the experimental data. To address cancer-related questions and generate interesting predictions, phenomenological properties such as survival, apoptosis, proliferation, EMT, metastasis and migration have been defined as model readouts. The linking between model readouts and the regulatory network has been set to reflect a particular activity status of the network components, which are reported as

markers for the correspondent phenomenological property. The usage of these readouts has been quite useful for model validation and testing predictions. Except for EMT and metastasis, the predicted phenomenological properties are relatively easy to access in cell culture with commercially available assays. For EMT and metastasis, commercially available cell adhesion and migration assays have been used as proxies, together with the quantification of E-cadherin as a marker of EMT [131]. Taken all together, the logical modelling framework has been proven to successfully identify and understand combinatorial effects of perturbations of regulatory networks in cancer.

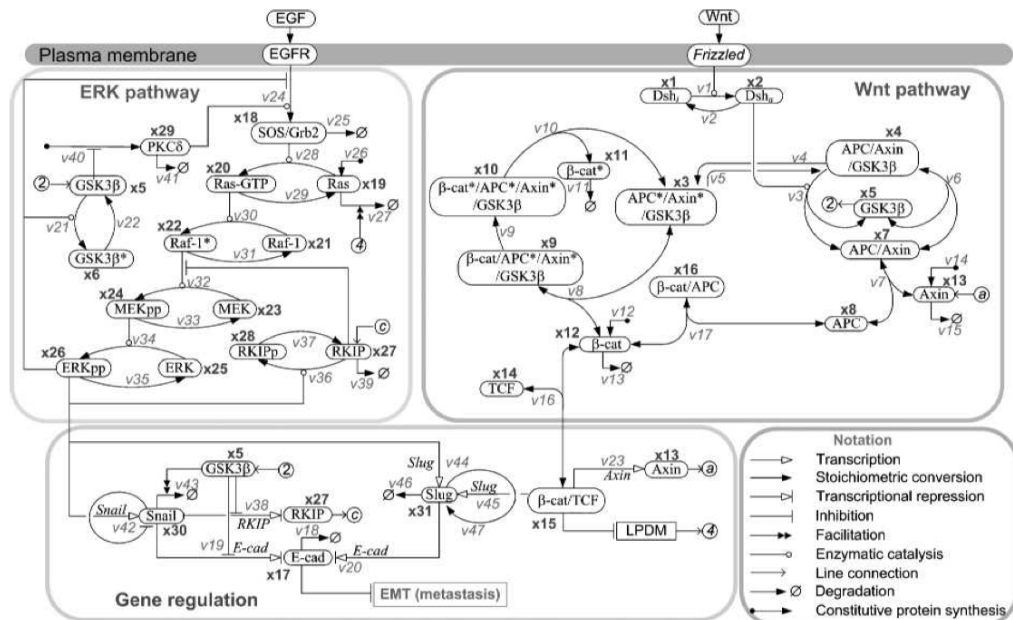
### **3.2. Mathematical models of EMT**

During the last decade, mathematical models have been developed to address particular questions about the regulation of EMT in the cancer context. Both continuous (kinetic modelling) and discrete (logical modelling) frameworks have been used as modelling approaches for describing EMT properties. In the following subsections, we will briefly describe these modelling works and their main objectives and findings.

#### **3.2.1 Kinetic models of EMT**

With the purpose of analysing the behaviours of multiple regulatory feedback loops between MAPK and Wnt signalling during EMT, Shin et al. developed an ODE-based continuous model that describes the expression of E-cadherin (EMT readout) in response to EGF and Wnt [8,20,151]. In this model (Figure 13), the inhibition of E-cadherin transcription by Snail1 and Snail2 was considered together with the mechanisms of Snail activation by Wnt and ERK [173]. The authors concluded: 1) Coupling positive feedback loops containing the inhibition of Raf kinase inhibitor protein (RKIP) by ERK with the transcriptional repression

of RKIP by Snail are essential to generate a switch-like behaviour of E-cadherin expression; 2) RKIP expression, inhibits EMT progression by preventing MAPK signalling and E-cadherin suppression [173].

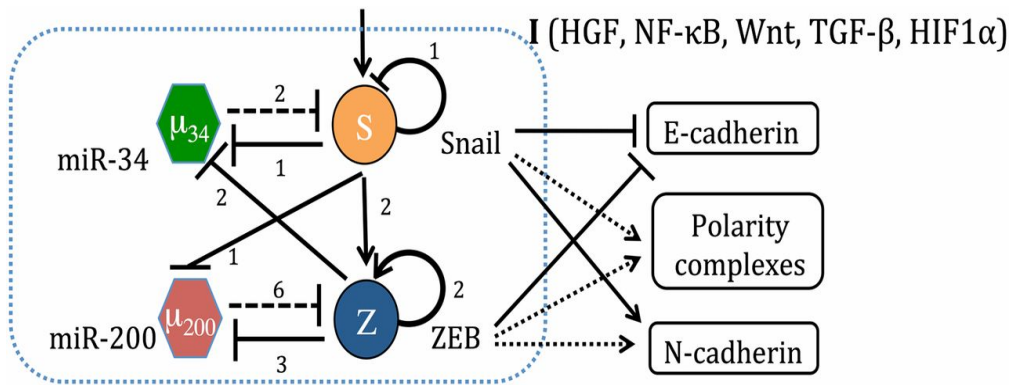


**Figure 13. Network representation of the mechanisms of EMT by EGF and Wnt signalling, modelled in the work of Shin et al.** The rates laws in the model are represented by the letter *v* and the variables in the model by the letter *x*. Figure reproduced from [173].

The mutual inhibitions between the EMT master transcription regulators (Zeb and Snail) and the microRNAs (miR200 and miR34 families) have been proposed to be the core regulatory circuit of EMT (Figure 14), and have been modelled through ODE in order to explain several observed cell fate properties of EMT [174–177]. In the work of M. Lu et al., the function of the miR200/Zeb and miR34/Snail modules of the regulatory EMT circuit was analysed [174]. The authors found that miR34/Snail circuit acts as a noise buffering integrator that depends on the strength of external and internal signals, controlling EMT initiation

[174]. On the other hand, ZEB/miR200 forms a circuit with 3 stable states; high/low, medium/medium and low/high levels, which were associated to Epithelial, Hybrid and Mesenchymal phenotypes [174].

In another work, the core regulatory circuit of EMT (Figure 14) was analysed to explain the reversible behaviour of EMT (MET) in response to TGF $\beta$  [175]. Through deterministic and stochastic analysis, the authors showed that the EMT circuit explains EMT and partial EMT in a two-step process (Epithelial  $\rightarrow$  Hybrid  $\rightarrow$  Mesenchymal), where the Hybrid state is an intermediate and depends on the strength and duration of the TGF $\beta$  signal [175]. In this work, they also found that the miR-34/Snail1 mutual inhibition was responsible for the reversible switch and regulates the initiation of EMT, whereas the ZEB/miR-200 controls the establishment of an irreversible Mesenchymal state by autocrine regulation of TGF $\beta$ , generating hysteresis behaviour [175]. Except for the autocrine regulation of TGF $\beta$ , the predictions from the modelling work of Lu et al. and Tian et al. were confirmed experimentally in cell cultures using MCDK cell lines [115]. Later on, detailed kinetic models accounting for the molecular mechanisms that could control the mutual inhibitions between ZEB1/2 and miR200b/c have further demonstrated the dose-dependent autocrine TGF $\beta$  effect on the stabilization of Mesenchymal phenotype [177]. These models were validated by comparing predictions with series of time course kinetics that followed the concentrations of miR200b, miR200b, Zeb1, Zeb2 an in MCDK cell lines exposed to different concentrations of TGF $\beta$  [177].

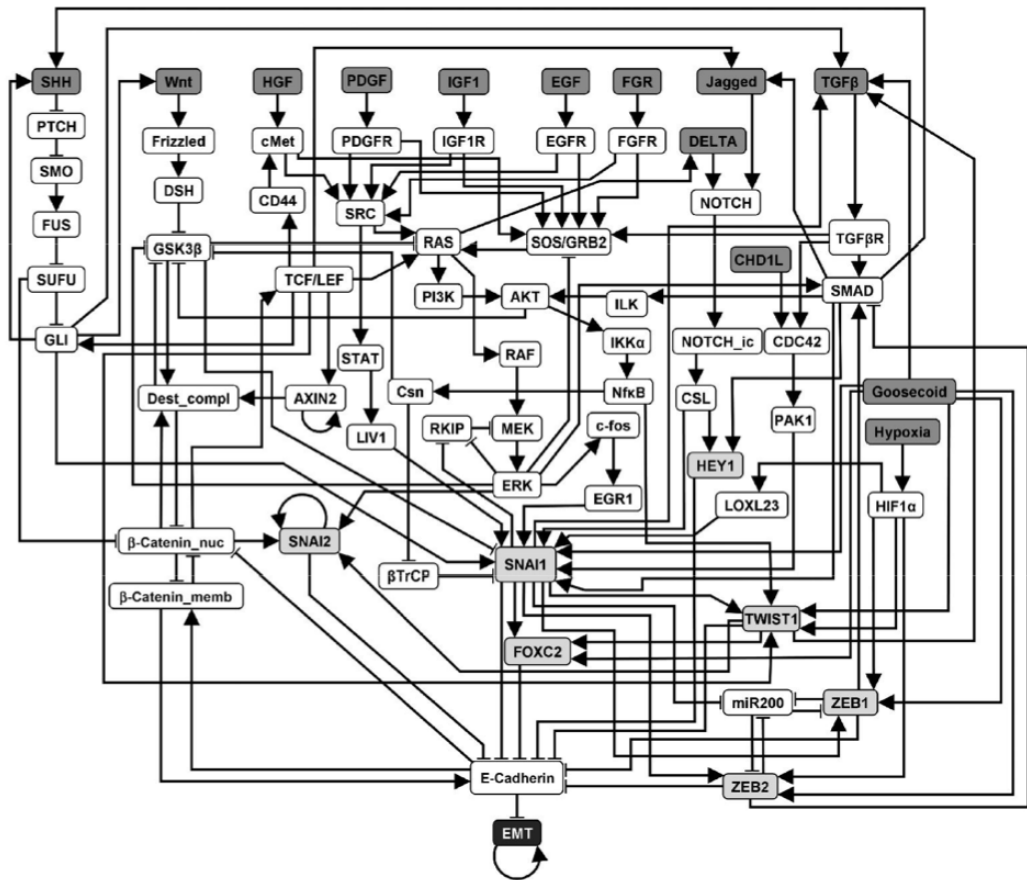


**Figure 14. The proposed core regulatory circuit of EMT.** The circuit is composed of two linked mutual inhibition circuits, the miR200/ZEB and the miR34/Snail. Activations are represented by arrows and inhibitions by a solid bar. Solid lines denote transcriptional mechanisms and dashed lines denote the translational regulation mechanisms triggered by microRNAs. Dotted lines indicate indirect regulation. The numbers along the arrows represent the number of binding sites involved, deduced from experimental data. Figure reproduced from [174].

The transitions between Mesenchymal, Hybrid and Amoeboid-like have been also analysed by coupling the core regulatory circuit of EMT with the mutual inhibitory circuit of Rac1/RhoA that determines the Mesenchymal-to-Amoeboid transition [35,176]. In this model, the authors predicted that the overall strength of the miRNAs inhibition on Snail/Zeb family is critical to determine the repression of EMT/MET, and eventually results in hybrid phenotypes [176]. This means that, at low degree of microRNAs, the model predicts that Amoeboid-like/Mesenchymal cell fate is favoured, dependent on the degree of activation of Gab1 and Grb2 co-receptors by HGF receptor [35]. On the other hand, strong microRNAs levels were predicted to stabilize the Hybrid phenotype [176]. Recently, Jolly et al. identified additional stability factors in lung cancer cell lines, that are required to be coupled with EMT regulatory core circuit in order to explain how miR200 can stabilize Hybrid phenotypes during carcinomas migration [19].

### 3.2.2. Logical models of EMT

A Boolean logical model was recently developed by Steinway et al. to understand how TGF $\beta$  drives EMT in hepatocarcinoma (Figure 15) [131]. So far, this model is the most detailed literature based EMT regulatory network, accounting for several signalling pathways and transcription factors reported to be involved in EMT of hepatocarcinomas [131]. The analysis of the model showed a joint activation of Wnt, Sonic hedgehog and TGF $\beta$  signalling as a conserved axis of the mesenchymal phenotype during EMT in hepatocytes. This prediction was experimentally confirmed *in vitro* using murine and hepatocarcinoma cell lines. Moreover, the model analysis predicted a total of 8 regulatory circuits involved in EMT and in the stabilization of the mesenchymal phenotype, including the Wnt/ $\beta$ -catenin feedback loop and the miR200/ZEB crosstalk with both MAPK/SMAD circuit and the TGF $\beta$  feedback loop. The model of Steinway et al. was further used in a posterior research work that explored all possible combinations of single, double and triple KO of model components [170]. This was done with the purpose of identifying effective drug combinations that could improve the usage of SMAD drug inhibitors in cancer therapy. In this work, the model predicted that SMAD inhibition in combination with NOTCH, RAS, AKT or PI3K would be more effective in preventing TGF $\beta$  driven EMT in hepatocarcinoma [170]. These predictions were tested in cell cultures of hepatocarcinoma cell lines by measuring E-cadherin expression and with comparing migration assays. In this work, several combinations of network perturbations, including SMAD inhibition, have resulted in stable states that were compatible with both Hybrid and Intermediate phenotypes (partial EMT phenotypes).



**Figure 15. Network model of EMT signalling in hepatocarcinoma cell lines.** The nodes represent molecular entities such as proteins, genes, miRNAs and other biological molecules. Nodes in white are regulatory components, dark gray represent upstream signals, nodes in light gray the downstream signals (transcription factors), and the node in black with white text the output of the model (EMT). Edges represent activation or inhibitory effects between nodes. Full names of the nodes, logical rules and references of the model are in [131]. Figure reproduced from [131].

Meanwhile, another logical network model was developed for predicting the role of individual and combinations of mutations driving the metastasis potential of EMT in carcinomas [171]. Additionally to EMT, the model accounted for migration, invasion, apoptosis and cell cycle arrest as phenomenological outputs [171]. The model was analysed considering generic ECM and DNA damage as two external



signals from the microenvironment. In this study, the model recapitulated experiments of TGF $\beta$  induced in lung carcinoma cell lines and reported mutations. The model analysis identified combinations of mutational effects (gene knockout and overexpression) that resulted in a synergistic increase of the probability of reaching the metastasis fate. Interestingly, the authors identified a synergism for the combination of NOTCH over-activation with p53 knockout leading to high cancer cell migration, which was in agreement with the results observed in a mouse model of gut cancer [171].

### **3.3. Modelling the effects of microenvironment signals on cell adhesion**

So far, how multiple microenvironment conditions affect Mesenchymal and Hybrid cell fates have not been analysed accounting for mutational perturbations in cancer. In addition, the role of cell-cell contact signals has not been yet explored in the context of carcinoma invasion and metastasis. Moreover, mathematical models that include features (e.g. cell-cell adhesion and focal adhesion) that characterize the distinction between invasive (e.g. Mesenchymal or Hybrid) and non-invasive phenotypes (e.g. Epithelial) have not been yet developed. Thus, the development of such models would provide a proxy to understand the effect of the microenvironment on cancer migration and metastasis acquisition in an holistic manner. This calls for new mathematical models to integrate the current knowledge and provide a framework that establishes a relation between microenvironment and phenotypes. In the next chapter of this thesis, we will present a novel logical model that accounts for the effects of 8 microenvironment signals on the regulations of cell-cell adhesion and focal adhesion. Moreover, in the next chapter, we will also analyse the model towards addressing the following main questions, previously highlighted during this chapter.

- 1) How does the microenvironment control the switching between Epithelial, Hybrid and Mesenchymal phenotypes?
- 2) What is the role of cell-cell contact signals on the switching between phenotypes?
- 3) How does mutational perturbations in cancer interfere with the switching between phenotypes?

#### 4. References

1. James SL, Peter GA. Epithelial Cells. In: Lowe's, Stevens, editors. Human Histology. Fourth Edi. Elsevier; 2015. pp. 37–38.
2. Leopold PL, Vincent J, Wang H. A comparison of epithelial-to-mesenchymal transition and re- epithelialization. *Semin Cancer Biol.* 2012;22: 471–483. doi:10.1016/j.semcancer.2012.07.003
3. Lee K, Nelson CM. New Insights into the Regulation of Epithelial-Mesenchymal Transition and Tissue Fibrosis. 1st ed. International Review of Cell and Molecular Biology. Elsevier Inc.; 2012. doi:10.1016/B978-0-12-394305-7.00004-5
4. Kalluri R, Weinberg R a. The basics of epithelial-mesenchymal transition. *J Clin Invest.* 2009;119: 1420–1428. doi:10.1172/JCI39104.1420
5. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. *J Mammary Gland Biol Neoplasia.* 2010;15: 117–34. doi:10.1007/s10911-010-9178-9
6. Savagner P. Epithelial-mesenchymal transitions: from cell plasticity to concept elasticity. *Curr Top Dev Biol.* 2015;112: 273–300. doi:10.1016/bs.ctdb.2014.11.021
7. Tsai JH, Yang J. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes Dev.* 2013;27: 2192–206. doi:10.1101/gad.225334.113
8. Nieto MA, Cano A. The epithelial-mesenchymal transition under control: Global programs to regulate epithelial plasticity. *Semin Cancer Biol.* Elsevier Ltd; 2012;22: 361–368. doi:10.1016/j.semcancer.2012.05.003
9. Larue L, Bellacosa A. Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene.* 2005;24: 7443–54. doi:10.1038/sj.onc.1209091
10. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* Nature Publishing Group; 2014;15: 178–96. doi:10.1038/nrm3758

11. Chui MH. Insights into cancer metastasis from a clinicopathologic perspective: Epithelial-Mesenchymal Transition is not a necessary step. *Int J Cancer*. 2013;132: 1487–95. doi:10.1002/ijc.27745
12. Zeisberg M, Neilson EG. Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest*. American Society for Clinical Investigation; 2009;119: 1429–37. doi:10.1172/JCI36183
13. Klymkowsky MW, Savagner P. Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. *Am J Pathol*. 2009;174: 1588–93. doi:10.2353/ajpath.2009.080545
14. Yang J, Weinberg R a. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell*. 2008;14: 818–29. doi:10.1016/j.devcel.2008.05.009
15. Steinestel K, Eder S, Schrader AJ, Steinestel J. Clinical significance of epithelial-mesenchymal transition. *Clin Transl Med*. 2014;3: 1–13.
16. Le Bras GF, Taubenslag KJ, Andl CD. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adh Migr*. Taylor & Francis; 2012;6: 365–73. doi:10.4161/cam.21326
17. Wieczorek E, Jablonska E, Wasowicz W, Reszka E. Matrix metalloproteinases and genetic mouse models in cancer research: a mini-review. *Tumour Biol*. 2014; 163–175. doi:10.1007/s13277-014-2747-6
18. Kim SH, Turnbull J, Guimond S. Extracellular matrix and cell signalling: The dynamic cooperation of integrin, proteoglycan and growth factor receptor. *J Endocrinol*. 2011;209: 139–151. doi:10.1530/JOE-10-0377
19. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget*. Impact Journals, LLC; 2016;7: 27067–27084. doi:10.18632/oncotarget.8166
20. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Onuchic JN, et al. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. 2015;5: 1–19. doi:10.3389/fonc.2015.00155
21. Garg M. Epithelial, mesenchymal and hybrid epithelial/mesenchymal phenotypes and their clinical relevance in cancer metastasis. *Expert Rev Mol Med*. 2017;19: e3. doi:10.1017/erm.2017.6
22. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Biol*. Nature Publishing Group; 2009;10: 445–457. doi:10.1038/nrm2720
23. Yamaguchi N, Mizutani T, Kawabata K, Haga H, Friedl P, Gilmour D, et al. Leader cells regulate collective cell migration via Rac activation in the downstream signaling of integrin  $\beta$ 1 and PI3K. *Sci Rep*. Nature Publishing Group; 2015;5: 7656. doi:10.1038/srep07656
24. Friedl P, Alexander S. Cancer invasion and the microenvironment: Plasticity and reciprocity. *Cell*. Elsevier Inc.; 2011;147: 992–1009. doi:10.1016/j.cell.2011.11.016

25. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin.* 2017;67: 7–30. doi:10.3322/caac.21387
26. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWW, Comber H, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. *Eur J Cancer.* 2013;49: 1374–1403. doi:10.1016/j.ejca.2012.12.027
27. van de Wouw A., Janssen-Heijnen ML., Coebergh JW., Hillen HF., Hemminki K, Fiorentini G, et al. Comparison of survival of patients with metastases from known versus unknown primaries: survival in metastatic cancer. *BMC Cancer. BioMed Central;* 2013; 13–36. doi:10.1016/S0959-8049(01)00378-1
28. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell.* 2011;144: 646–674. doi:10.1016/j.cell.2011.02.013
29. Floor S, van Staveren WCG, Larsimont D, Dumont JE, Maenhaut C. Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating-cancer stem cells: distinct, overlapping or same populations. *Oncogene.* 2011;30: 4609–21. doi:10.1038/onc.2011.184
30. Clark AG, Vignjevic DM. Modes of cancer cell invasion and the role of the microenvironment. *Curr Opin Cell Biol.* 2015;36: 13–22. doi:10.1016/j.ceb.2015.06.004
31. Kawauchi T. Cell adhesion and its endocytic regulation in cell migration during neural development and cancer metastasis. *Int J Mol Sci. Multidisciplinary Digital Publishing Institute (MDPI);* 2012;13: 4564–90. doi:10.3390/ijms13044564
32. Hai R, García-Sastre A, Swayne DE, Palese P. A reassortment-incompetent live attenuated influenza virus vaccine for protection against pandemic virus strains. *J Virol.* 2011;85: 6832–43. doi:10.1128/JVI.00609-11
33. Krakhmal N V, Zavyalova M V, Denisov E V, Vtorushin S V, Perelmuter VM. Cancer invasion: Patterns and mechanisms [Internet]. *Acta Naturae. Park Media;* 2015. pp. 17–28. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26085941>
34. Panková K, Rösel D, Novotný M, Brábek J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cell Mol Life Sci. Springer;* 2010;67: 63–71. doi:10.1007/s00018-009-0132-1
35. Huang B, Lu M, Jolly MK, Tsarfaty I, Onuchic J, Ben-Jacob E. The three-way switch operation of Rac1/RhoA GTPase-based circuit controlling amoeboid-hybrid-mesenchymal transition. *Sci Rep.* 2014;4: 6449. doi:10.1038/srep06449
36. van Zijl F, Krupitza G, Mikulits W. Initial steps of metastasis: Cell invasion and endothelial transmigration. *Mutat Res Mutat Res.* 2011;728: 23–34. doi:10.1016/j.mrrev.2011.05.002
37. Murphy DA, Courtneidge SA. The ‘ins’ and ‘outs’ of podosomes and invadopodia: characteristics, formation and function. *Nat Rev Mol Cell Biol.* 2011;12: 413–426. doi:10.1038/nrm3141
38. Tsuji T, Ibaragi S, Hu G. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res.* 2009;69: 7135–9. doi:10.1158/0008-5472.CAN-09-1618

39. Canel M, Serrels A, Frame MC, Brunton VG. E-cadherin–integrin crosstalk in cancer invasion and metastasis. *J Cell Sci.* 2013;126: 393–401. doi:10.1242/jcs.100115
40. Nagano M, Hoshino D, Koshikawa N, Akizawa T, Seiki M. Turnover of focal adhesions and cancer cell migration. *Int J Cell Biol.* 2011;2012: 1–12. doi:10.1155/2012/310616
41. Kim D-H, Wirtz D. Predicting how cells spread and migrate: focal adhesion size does matter. *Cell Adh Migr.* Taylor & Francis; 2013;7: 293–6. doi:10.4161/cam.24804
42. Wu C. Focal adhesion: a focal point in current cell biology and molecular medicine. *Cell Adh Migr.* Taylor & Francis; 2007;1: 13–8. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19262093>
43. Winograd-Katz SE, Fässler R, Geiger B, Legate KR. The integrin adhesome: from genes and proteins to human disease. *Nat Rev Mol Cell Biol.* 2014;15: 273–288. doi:10.1038/nrm3769
44. Zaidel-Bar R, Itzkovitz S, Ma'ayan A, Iyengar R, Geiger B. Functional atlas of the integrin adhesome. *Cell.* 2009;9: 858–867. doi:10.1038/ncb0807-858.Functional
45. Vicente-Manzanares M, Horwitz AR. Adhesion dynamics at a glance. *J Cell Sci.* Company of Biologists; 2011;124: 3923–3927. doi:10.1242/jcs.095653
46. Legate KR, Wickström S a, Fässler R, Fa R, Wickstro S a. Genetic and cell biological analysis of integrin outside-in signaling. *Genes Dev.* 2009;23: 397–418. doi:10.1101/gad.1758709
47. Eberwein P, Laird D, Schulz S, Reinhard T, Steinberg T, Tomakidi P. Modulation of focal adhesion constituents and their down-stream events by EGF: On the cross-talk of integrins and growth factor receptors. *Biochim Biophys Acta - Mol Cell Res.* 2015;1853: 2183–2198. doi:10.1016/j.bbamcr.2015.06.004
48. Ricono JM, Huang M, Barnes LA, Lau SK, Weis SM, Schlaepfer DD, et al. Specific cross-talk between epidermal growth factor receptor and integrin alphavbeta5 promotes carcinoma cell invasion and metastasis. *Cancer Res.* 2009;69: 1383–91. doi:10.1158/0008-5472.CAN-08-3612
49. Chan PM, Lim L, Manser E. PAK Is Regulated by PI3K, PIX, CDC42, and PP2C and Mediates Focal Adhesion Turnover in the Hyperosmotic Stress-induced p38 Pathway. *J Biol Chem.* 2008;283: 24949–24961. doi:10.1074/jbc.M801728200
50. Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, et al. Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell.* 2002;108: 837–47. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11955436>
51. Mitra SK, Schlaepfer DD. Integrin-regulated FAK–Src signaling in normal and cancer cells. *Curr Opin Cell Biol.* 2006;18: 516–523. doi:10.1016/j.ceb.2006.08.011
52. Bolós V, Gasent JM, López-Tarruella S, Grande E. The dual kinase complex FAK–Src as a promising therapeutic target in cancer. *Onco Targets Ther.* 2010;3: 83–97. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20616959>

53. Schaller MD, Hildebrand JD, Parsons JT. Complex formation with focal adhesion kinase: A mechanism to regulate activity and subcellular localization of Src kinases. *Mol Biol Cell*. American Society for Cell Biology; 1999;10: 3489–505. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10512882>
54. Mierke CT, Frey B, Fellner M, Herrmann M, Fabry B. Integrin  $\alpha 5 \beta 1$  facilitates cancer cell invasion through enhanced contractile forces. *J Cell Sci*. Company of Biologists; 2011;124: 369–83. doi:10.1242/jcs.071985
55. Humphries JD, Byron A, Humphries MJ. Integrin ligands at a glance. *J Cell Sci*. 2006;119: 3901–3903. doi:10.1242/jcs.03098
56. Gumbiner BM. Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol*. 2005;6: 622–34. doi:10.1038/nrm1699
57. Hartsock A, Nelson WJ. Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta - Biomembr*. NIH Public Access; 2008;1778: 660–669. doi:10.1016/j.bbmem.2007.07.012
58. Kowalczyk AP, Nanes BA. Adherens junction turnover: regulating adhesion through cadherin endocytosis, degradation, and recycling. *Subcell Biochem*. NIH Public Access; 2012;60: 197–222. doi:10.1007/978-94-007-4186-7\_9
59. Adams CL, Chen YT, Smith SJ, Nelson WJ. Mechanisms of epithelial cell-cell adhesion and cell compaction revealed by high-resolution tracking of E-cadherin-green fluorescent protein. *J Cell Biol*. 1998;142: 1105–19. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9722621>
60. Bryant DM, Stow JL. The ins and outs of E-cadherin trafficking. *Trends Cell Biol*. 2004;14: 427–34. doi:10.1016/j.tcb.2004.07.007
61. Ireton RCRC, Davis MA, van Hengel J, Mariner DJ, Barnes K, Thoreson MA, et al. A novel role for p120 catenin in E-cadherin function. *J cell Biol*. 2002;159: 465–476. doi:10.1083/jcb.200205115
62. Davis M a., Ireton RC, Reynolds AB. A core function for p120-catenin in cadherin turnover. *J Cell Biol*. 2003;163: 525–534. doi:10.1083/jcb.200307111
63. Nelson WJ. Regulation of cell-cell adhesion by the cadherin-catenin complex. *Biochem Soc Trans*. NIH Public Access; 2008;36: 149–55. doi:10.1042/BST0360149
64. Chen YT, Stewart DB, Nelson WJ. Coupling assembly of the E-cadherin/beta-catenin complex to efficient endoplasmic reticulum exit and basal-lateral membrane targeting of E-cadherin in polarized MDCK cells. *J Cell Biol*. 1999;144: 687–99. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10037790>
65. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the Cadherin-Catenin-Actin Complex. *Cell*. 2005;123: 889–901. doi:10.1016/j.cell.2005.09.020

66. Vasioukhin V, Bauer C, Yin M, Fuchs E. Directed actin polymerization is the driving force for epithelial cell-cell adhesion. *Cell*. 2000;100: 209–19. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10660044>
67. Daugherty RL, Gottardi CJ. Phospho-regulation of  $\beta$ -Catenin Adhesion and Signaling Functions. *Physiology*. 2008;3: 303–309. doi:10.1152/physiol.00020.2007
68. Shibamoto S, Hayakawa M, Takeuchi K, Hori T, Miyazawa K, Kitamura N, et al. Association of p120, a tyrosine kinase substrate, with E-cadherin/catenin complexes. *J Cell Biol*. 1995;128: 949–57. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7876318>
69. Roura S, Miravet S, Piedra J, García de Herreros A, Duñach M. Regulation of E-cadherin/Catenin association by tyrosine phosphorylation. *J Biol Chem*. 1999;274: 36734–40. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10593980>
70. Xu Y, Fisher GJ. Receptor type protein tyrosine phosphatases (RPTPs) - roles in signal transduction and human disease. *J Cell Commun Signal*. Springer; 2012;6: 125–38. doi:10.1007/s12079-012-0171-5
71. Ostman A, Böhmer FD. Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatases. *Trends Cell Biol*. 2001;11: 258–66. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11356362>
72. Pacquelet A, Lin L, Rorth P. Binding site for p120/beta-catenin is not required for Drosophila E-cadherin function in vivo. *J Cell Biol*. 2003;160: 313–319. doi:10.1083/jcb.200207160
73. Brücher BLDM, Jamall IS. Cell-Cell Communication in the Tumor Microenvironment, Carcinogenesis, and Anticancer Treatment. *Cell Physiol Biochem*. 2014;34: 213–243. doi:10.1159/000362978
74. Gao D, Vahdat LT, Wong S, Chang JC, Mittal V. Microenvironmental regulation of epithelial-mesenchymal transitions in cancer. *Cancer Res*. NIH Public Access; 2012;72: 4883–9. doi:10.1158/0008-5472.CAN-12-1223
75. Li H, Xu F, Li S, Zhong A, Meng X, Lai M. The tumor microenvironment: An irreplaceable element of tumor budding and epithelial-mesenchymal transition-mediated cancer metastasis. *Cell Adh Migr*. Taylor & Francis; 2016;10: 434–446. doi:10.1080/19336918.2015.1129481
76. Jing Y, Han Z, Zhang S, Liu Y, Wei L. Epithelial-Mesenchymal Transition in tumor microenvironment. *Cell Biosci*. BioMed Central; 2011;1: 29. doi:10.1186/2045-3701-1-29
77. Liu R-Y, Zeng Y, Lei Z, Wang L, Yang H, Liu Z, et al. JAK/STAT3 signaling is required for TGF- $\beta$ -induced epithelial-mesenchymal transition in lung cancer cells. *Int J Oncol*. 2014;44: 1643–51. doi:10.3892/ijo.2014.2310
78. Lo H-W, Hsu S-C, Xia W, Cao X, Shih J-Y, Wei Y, et al. Epidermal Growth Factor Receptor Cooperates with Signal Transducer and Activator of Transcription 3 to Induce Epithelial-Mesenchymal Transition in Cancer Cells via Up-regulation of TWIST Gene Expression. *Cancer Res*. 2007;67: 9066–9076. doi:10.1158/0008-5472.CAN-07-0575

79. Grotegut S, von Schweinitz D, Christofori G, Lehenbre FF. Hepatocyte growth factor induces cell scattering through MAPK/Egr-1-mediated upregulation of Snail. *EMBO J.* 2006;25: 3534–3545. doi:10.1038/sj.emboj.7601213
80. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. *Breast Cancer Res. BioMed Central*; 2011;13: 227. doi:10.1186/bcr2912
81. Landskron G, De la Fuente M, Thuwajit P, Thuwajit C, Hermoso MA. Chronic inflammation and cytokines in the tumor microenvironment. *J Immunol Res.* 2014;2014: 149185. doi:10.1155/2014/149185
82. Nakamura T, Sakai K, Nakamura T, Matsumoto K. Hepatocyte growth factor twenty years on: Much more than a growth factor. *J Gastroenterol Hepatol.* 2011;26: 188–202. doi:10.1111/j.1440-1746.2010.06549.x
83. Hardy KM, Booth BW, Hendrix MJC, Salomon DS, Strizzi L. ErbB/EGF Signaling and EMT in Mammary Development and Breast Cancer. *J Mammary Gland Biol Neoplasia.* 2010;15: 191–1.
84. Gilkes DM, Semenza GL, Wirtz D. Hypoxia and the extracellular matrix: drivers of tumour metastasis. *Nat Rev Cancer.* 2014;14: 430–439. doi:10.1038/nrc3726
85. Cheng JC, Leung PCK. Type I collagen down-regulates E-cadherin expression by increasing PI3KCA in cancer cells. *Cancer Lett. Elsevier Ireland Ltd*; 2011;304: 107–116. doi:10.1016/j.canlet.2011.02.008
86. Chaudhuri O, Koshy ST, Branco da Cunha C, Shin J-W, Verbeke CS, Allison KH, et al. Extracellular matrix stiffness and composition jointly regulate the induction of malignant phenotypes in mammary epithelium. *Nat Mater.* 2014;13: 970–8. doi:10.1038/nmat4009
87. Fujii K, Furukawa F, Matsuyoshi N. Ligand activation of overexpressed epidermal growth factor receptor results in colony dissociation and disturbed E-cadherin function in HSC-1 human cutaneous squamous carcinoma cells. *Exp Cell Res.* 1996;223: 50–62. doi:10.1006/excr.1996.0057
88. Tacchini L, De Ponti C, Matteucci E, Follis R, Desiderio MA. Hepatocyte growth factor-activated NF-kappaB regulates HIF-1 activity and ODC expression, implicated in survival, differently in different carcinoma cell lines. *Carcinogenesis.* 2004;25: 2089–100. doi:10.1093/carcin/bgh227
89. Pawlus MR, Wang L, Hu C-J. STAT3 and HIF1 $\alpha$  cooperatively activate HIF1 target genes in MDA-MB-231 and RCC4 cells. *Oncogene.* 2014;33: 1670–9. doi:10.1038/onc.2013.115
90. Xie G, Yao Q, Liu Y, Du S, Liu A, Guo Z, et al. IL-6-induced epithelial-mesenchymal transition promotes the generation of breast cancer stem-like cells analogous to mammosphere cultures. *Int J Oncol.* 2012;40: 1171–9. doi:10.3892/ijo.2011.1275
91. Cervellati F, Cervellati C, Romani A, Cremonini E, Sticozzi C, Belmonte G, et al. Hypoxia induces cell damage via oxidative stress in retinal epithelial cells. *Free Radic Res.* 2014;48: 303–12. doi:10.3109/10715762.2013.867484



92. Chen J, Imanaka N, Chen J, Griffin JD. Hypoxia potentiates Notch signaling in breast cancer leading to decreased E-cadherin expression and increased cell migration and invasion. *Br J Cancer*. 2010;102: 351–60. doi:10.1038/sj.bjc.6605486
93. Sahlgren C, Gustafsson M V., Jin S, Poellinger L, Lendahl U. Notch signaling mediates hypoxia-induced tumor cell migration and invasion. *Proc Natl Acad Sci*. 2008;105: 6392–6397. doi:10.1073/pnas.0802047105
94. Espinoza I, Miele L. Deadly crosstalk: Notch signaling at the intersection of EMT and cancer stem cells. *Cancer Lett*. 2013;341: 41–45. doi:10.1016/j.canlet.2013.08.027
95. Ding S, Zhang W, Xu Z, Xing C, Xie H, Guo H, et al. Induction of an EMT-like transformation and MET in vitro. *J Transl Med. BioMed Central*; 2013;11: 164. doi:10.1186/1479-5876-11-164
96. Deneff C. Contact-dependent Signaling. *Cell Commun Insights*. 2014; 1–11. doi:10.4137/CCi.s12484.TYPE
97. Katoh M. Function and cancer genomics of FAT family genes (Review). *Int J Oncol*. 2012;41: 1913–8. doi:10.3892/ijo.2012.1669
98. Mao Y, Francis-West P, Irvine KD. Fat4/Dchs1 signaling between stromal and cap mesenchyme cells influences nephrogenesis and ureteric bud branching. *Development*. 2015;142: 2574–85. doi:10.1242/dev.122630
99. Mohebiany AN, Nikolaienko RM, Bouyain S, Harroch S. Receptor-type tyrosine phosphatase ligands: looking for the needle in the haystack. *FEBS J*. 2013;280: 388–400. doi:10.1111/j.1742-4658.2012.08653.x
100. Sap J, Jiang YP, Friedlander D, Grumet M, Schlessinger J. Receptor tyrosine phosphatase R-PTP-kappa mediates homophilic binding. *Mol Cell Biol. American Society for Microbiology (ASM)*; 1994;14: 1–9. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8264577>
101. Ostman A, Yang Q, Tonks NK. Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhanced with increasing cell density. *Proc Natl Acad Sci U S A. National Academy of Sciences*; 1994;91: 9680–4. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7937872>
102. De Craene B, Berx G. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer. Nature Publishing Group*; 2013;13: 97–110. doi:10.1038/nrc3447
103. Waldmann J, Slater EP, Langer P, Buchholz M, Ramaswamy A, Walz MK, et al. Expression of the Transcription Factor Snail and Its Target Gene Twist Are Associated with Malignancy in Pheochromocytomas. *Ann Surg Oncol*. 2009;16: 1997–2005. doi:10.1245/s10434-009-0480-y
104. Yang Z, Rayala S, Nguyen D, Vadlamudi RK, Chen S, Kumar R. Pak1 phosphorylation of snail, a master regulator of epithelial-to-mesenchyme transition, modulates snail's subcellular localization and functions. *Cancer Res*. 2005;65: 3179–84. doi:10.1158/0008-5472.CAN-04-3480

105. de Herreros AG, Peiró S, Nassour M, Savagner P. Snail Family Regulation and Epithelial Mesenchymal Transitions in Breast Cancer Progression. *J Mammary Gland Biol Neoplasia*. 2010;15: 135–147. doi:10.1007/s10911-010-9179-8
106. Bolós V, Peinado H, Pérez-Moreno MA, Fraga MF, Esteller M, Cano A. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci*. 2003;116: 499–511. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12508111>
107. Medici D, Hay ED, Olsen BR. Snail and Slug Promote Epithelial-Mesenchymal Transition through  $\beta$ -Catenin-T-Cell Factor-4-dependent Expression of Transforming Growth Factor- $\beta$ . *Mol Biol Cell*. 2008;19: 4875–4887. doi:10.1091/mbc.E08-05-0506
108. Shinozaki A, Sakatani T, Ushiku T, Hino R, Isogai M, Ishikawa S, et al. Downregulation of MicroRNA-200 in EBV-Associated Gastric Carcinoma. *Cancer Res*. 2010;70. Available: <http://cancerres.aacrjournals.org/content/70/11/4719>
109. Hennessy BT, Gonzalez-Angulo A-M, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee J-S, et al. Characterization of a Naturally Occurring Breast Cancer Subset Enriched in Epithelial-to-Mesenchymal Transition and Stem Cell Characteristics. *Cancer Res*. 2009;69: 4116–4124. doi:10.1158/0008-5472.CAN-08-3441
110. Humphries B, Yang C. The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. *Oncotarget*. 2015;6: 6472–6498. doi:10.18632/oncotarget.3052
111. Xiong M, Jiang L, Zhou Y, Qiu W, Fang L, Tan R, et al. The miR-200 family regulates TGF- $\beta$ 1-induced renal tubular epithelial to mesenchymal transition through Smad pathway by targeting ZEB1 and ZEB2 expression. *AJP Ren Physiol*. 2012;302: F369–F379. doi:10.1152/ajprenal.00268.2011
112. Gill JG, Langer EM, Lindsley RC, Cai M, Murphy TL, Kyba M, et al. Snail and the microRNA-200 Family Act in Opposition to Regulate Epithelial-to-Mesenchymal Transition and Germ Layer Fate Restriction in Differentiating ESCs. *Stem Cells*. 2011;29: 764–776. doi:10.1002/stem.628
113. Bracken CP, Gregory PA, Kolesnikoff N, Bert AG, Wang J, Shannon MF, et al. A Double-Negative Feedback Loop between ZEB1-SIP1 and the microRNA-200 Family Regulates Epithelial-Mesenchymal Transition. *Cancer Res*. 2008;68: 7846–7854. doi:10.1158/0008-5472.CAN-08-1942
114. Valcourt U, Kowanzetz M, Niimi H, Heldin C-H, Moustakas A. TGF- $\beta$  and the Smad Signaling Pathway Support Transcriptomic Reprogramming during Epithelial-Mesenchymal Cell Transition. *Mol Biol Cell*. 2005;16: 1987–2002. doi:10.1091/mbc.E04-08-0658
115. Zhang J, Tian X-JX-J, Zhang H, Teng Y, Li R, Bai F, et al. TGF $\beta$ 1-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci Signal*. 2014;7: ra91-ra91. doi:10.1126/scisignal.2005304

116. Wang SE, Wu FY, Shin I, Qu S, Arteaga CL. Transforming growth factor {beta} (TGF-{beta})-Smad target gene protein tyrosine phosphatase receptor type kappa is required for TGF-{beta} function. *Mol Cell Biol.* American Society for Microbiology (ASM); 2005;25: 4703–15. doi:10.1128/MCB.25.11.4703-4715.2005
117. Piek E, Moustakas A, Kurisaki A, Heldin CH, ten Dijke P. TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. *J Cell Sci.* 1999;112 ( Pt 24): 4557–68. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10574705>
118. Mayer IA, Arteaga CL. The PI3K/AKT Pathway as a Target for Cancer Treatment. *Annu Rev Med.* 2015;67: 11–28. doi:10.1146/annurev-med-062913-051343
119. Bachelder RE, Yoon S-O, Franci C, de Herreros AG, Mercurio AM. Glycogen synthase kinase-3 is an endogenous inhibitor of Snail transcription: implications for the epithelial-mesenchymal transition. *J Cell Biol.* 2005;168: 29–33. doi:10.1083/jcb.200409067
120. Zhou BP, Deng J, Xia W, Xu J, Li YM, Gunduz M, et al. Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition. *Nat Cell Biol.* 2004;6: 931–40. doi:10.1038/ncb1173
121. Julien S, Puig I, Caretti E, Bonaventure J, Nelles L, van Roy F, et al. Activation of NF-κB by Akt upregulates Snail expression and induces epithelium mesenchyme transition. *Oncogene.* 2007;26: 7445–7456. doi:10.1038/sj.onc.1210546
122. Lemieux E, Bergeron SS, Durand VV, Asselin C, Saucier C, Rivard N. Constitutively active MEK1 is sufficient to induce epithelial-to-mesenchymal transition in intestinal epithelial cells and to promote tumor invasion and metastasis. *Int J Cancer.* 2009;125: 1575–1586. doi:10.1002/ijc.24485
123. Conacci-Sorrell M, Simcha I, Ben-Yedidia T, Blechman J, Savagner P, Ben-Ze'ev A. Autoregulation of E-cadherin expression by cadherin?cadherin interactions. *J Cell Biol.* 2003;163: 847–857. doi:10.1083/jcb.200308162
124. Farahani E, Patra HK, Jangamreddy JR, Rashedi I, Kawalec M, Rao Pariti RK, et al. Cell adhesion molecules and their relation to (cancer) cell stemness. *Carcinogenesis.* 2014;35: 747–759. doi:10.1093/carcin/bgu045
125. Bill HM, Knudsen B, Moores SL, Muthuswamy SK, Rao VR, Brugge JS, et al. Epidermal growth factor receptor-dependent regulation of integrin-mediated signaling and cell cycle entry in epithelial cells. *Mol Cell Biol.* 2004;24: 8586–99. doi:10.1128/MCB.24.19.8586-8599.2004
126. Cruet-Hennequart S, Maubant S, Luis J, Gauduchon P, Staedel C, Dedhar S. alpha(v) integrins regulate cell proliferation through integrin-linked kinase (ILK) in ovarian cancer cells. *Oncogene.* 2003;22: 1688–702. doi:10.1038/sj.onc.1206347
127. Peroukides S, Bravou V, Varakis J, Alexopoulos A, Kalofonos H, Papadaki H. ILK overexpression in human hepatocellular carcinoma and liver cirrhosis correlates with

- activation of Akt. *Oncol Rep.* 2008;20: 1337–44. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19020711>
128. Yuan Y, Xiao Y, Li Q, Liu Z, Zhang X, Qin C, et al. In vitro and in vivo effects of short hairpin RNA targeting integrin-linked kinase in prostate cancer cells. *Mol Med Rep.* 2013;8: 419–24. doi:10.3892/mmr.2013.1532
  129. Santibañez JF. JNK mediates TGF-beta1-induced epithelial mesenchymal transdifferentiation of mouse transformed keratinocytes. *FEBS Lett.* 2006;580: 5385–5391. doi:10.1016/j.febslet.2006.09.003
  130. Howard S, Deroo T, Fujita Y, Itasaki N. A positive role of cadherin in Wnt/ $\beta$ -catenin signalling during epithelial-mesenchymal transition. *PLoS One.* 2011;6: e23899. doi:10.1371/journal.pone.0023899
  131. Steinway SN, Zanudo JGT, Ding W, Rountree CB, Feith DJ, Loughran TP, et al. Network modeling of TGF $\beta$  signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint Sonic hedgehog and Wnt pathway activation. *Cancer Res.* 2014;74: 5963–5977. doi:10.1158/0008-5472.CAN-14-0225
  132. MacDonald BT, Tamai K, He X. Wnt/ $\beta$ -Catenin Signaling: Components, Mechanisms, and Diseases. *Dev Cell.* 2009;17: 9–26. doi:10.1016/j.devcel.2009.06.016
  133. Eastman Q, Grosschedl R. Regulation of LEF-1/TCF transcription factors by Wnt and other signals. *Curr Opin Cell Biol.* 1999;11: 233–240. doi:10.1016/S0955-0674(99)80031-3
  134. Alcaraz A, Mrowiec A, Insausti CL, García-Vizcaíno EM, Ruiz-Canada C, López-Martínez MC, et al. Autocrine TGF- $\beta$  Induces Epithelial to Mesenchymal Transition in Human Amniotic Epithelial Cells. *Cell Transplant.* 2013;22: 1351–1367. doi:10.3727/096368912X657387
  135. Szeto SG, Narimatsu M, Lu M, He X, Sidiqi AM, Tolosa MF, et al. YAP/TAZ Are Mechanoregulators of TGF- $\beta$ -Smad Signaling and Renal Fibrogenesis. *J Am Soc Nephrol.* 2016;27: 3117–3128. doi:10.1681/ASN.2015050499
  136. Varelas X, Miller BW, Sopko R, Song S, Gregorieff A, Fellouse FA, et al. The Hippo pathway regulates Wnt/beta-catenin signaling. *Dev Cell.* 2010;18: 579–91. doi:10.1016/j.devcel.2010.03.007
  137. Kim M, Jho E-H. Cross-talk between Wnt/ $\beta$ -catenin and Hippo signaling pathways: a brief review. *BMB Rep.* 2014;47: 540–5. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25154721>
  138. Garcia R, Bowman TL, Niu G, Yu H, Minton S, Muro-Cacho CA, et al. Constitutive activation of Stat3 by the Src and JAK tyrosine kinases participates in growth regulation of human breast carcinoma cells. *Oncogene.* 2001;20. doi:10.1038/sj.onc.1204349
  139. Shao S, Zhao XX, Zhang X, Luo M, Zuo X, Huang S, et al. Notch1 signaling regulates the epithelial-mesenchymal transition and invasion of breast cancer in a Slug-dependent manner. *Mol Cancer.* 2015;3: 14–28. doi:10.1186/s12943-015-0295-3

140. Purow BW, Sundaresan TK, Burdick MJ, Kefas BA, Comeau LD, Hawkinson MP, et al. Notch-1 regulates transcription of the epidermal growth factor receptor through p53. *Carcinogenesis*. 2008;29: 918–925. doi:10.1093/carcin/bgn079
141. Espinosa L, Cathelin S, D'Altri T, Trimarchi T, Statnikov A, Guiu J, et al. The Notch/Hes1 Pathway Sustains NF- $\kappa$ B Activation through CYLD Repression in T Cell Leukemia. *Cancer Cell*. 2010;18: 268–281. doi:10.1016/j.ccr.2010.08.006
142. Unno J, Satoh K, Hirota M, Kanno A, Hamada S, Ito H, et al. LIV-1 enhances the aggressive phenotype through the induction of epithelial to mesenchymal transition in human pancreatic carcinoma cells. *Int J Oncol*. 2009;35: 813–21. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19724917>
143. Noble D. *The music of life: Biology beyond the genome*. Oxford: Oxford University Press; 2006.
144. von Bertalanffy L. *General System theory: Foundations, Development, Applications*. 28 March 1. George Braziller, editor. 1968.
145. Breitling R. What is systems biology? *Front Physiol*. Frontiers Media SA; 2010;1: 9. doi:10.3389/fphys.2010.00009
146. Mesarovic MD. *Systems Theory and Biology*. Berlin: Springer-Verlag; 1968.
147. Westerhoff H V., Winder C, Messiha H, Simeonidis E, Adamczyk M, Verma M, et al. Systems Biology: The elements and principles of Life. *FEBS Lett*. 2009;583: 3882–3890. doi:10.1016/j.febslet.2009.11.018
148. Chuang H-Y, Hofree M, Ideker T. A decade of systems biology. *Annu Rev Cell Dev Biol*. NIH Public Access; 2010;26: 721–44. doi:10.1146/annurev-cellbio-100109-104122
149. Somvanshi PR, Venkatesh K V. A conceptual review on systems biology in health and diseases: from biological networks to modern therapeutics. *Syst Synth Biol*. Springer; 2014;8: 99–116. doi:10.1007/s11693-013-9125-3
150. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502: 333–9. doi:10.1038/nature12634
151. Kuperstein I, Bonnet E, Nguyen H-AH, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. *Oncogenesis*. 2015;4: 1–14. doi:10.1038/oncsis.2015.19
152. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. Oxford University Press; 2017;45: D353–D361. doi:10.1093/nar/gkw1092
153. Kuperstein I, Cohen DPA, Pook S, Viara E, Calzone L, Barillot E, et al. NaviCell: a web-based environment for navigation, curation and maintenance of large molecular interaction maps. *BMC Syst Biol*. 2013;7: 100. doi:10.1186/1752-0509-7-100

154. Kholodenko B, Yaffe MB, Kolch W. Computational approaches for analyzing information flow in biological networks. *Sci Signal*. 2012;5: 1–14. doi:10.1126/scisignal.2002961
155. Le Novère N. Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet*. Europe PMC Funders; 2015;16: 146–58. doi:10.1038/nrg3885
156. Sugita M. Functional analysis of chemical systems in vivo using a logical circuit equivalent. II. The idea of a molecular automaton. *J Theor Biol*. Academic Press; 1963;4: 179–192. doi:10.1016/0022-5193(63)90027-4
157. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol*. Academic Press; 1969;22: 437–467. doi:10.1016/0022-5193(69)90015-0
158. Thomas R. Boolean formalization of genetic control circuits. *J Theor Biol*. 1973;42: 563–85. Available: <http://www.ncbi.nlm.nih.gov/pubmed/4588055>
159. Wynn ML, Consul N, Merajver SD, Schnell S. Logic-based models in systems biology: a predictive and parameter-free network analysis method. *Integr Biol (Camb)*. NIH Public Access; 2012;4: 1323–37. doi:10.1039/c2ib20193c
160. Abou-Jaoudé W, Traynard P, Monteiro PT, Saez-Rodriguez J, Helikar T, Thieffry D, et al. Logical Modeling and Dynamical Analysis of Cellular Networks. *Front Genet*. Frontiers Media SA; 2016;7: 94. doi:10.3389/fgene.2016.00094
161. Thomas R. Regulatory networks seen as asynchronous automata: A logical description. *J Theor Biol*. Academic Press; 1991;153: 1–23. doi:10.1016/S0022-5193(05)80350-9
162. Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger D a. Logic-based models for the analysis of cell signaling networks. *Biochemistry*. 2010;49: 3216–24. doi:10.1021/bi902202q
163. Bérenguier D, Chaouiya C, Monteiro PT, Naldi a, Remy E, Thieffry D, et al. Dynamical modeling and analysis of large cellular regulatory networks. *Chaos*. 2013;23: 025114. doi:10.1063/1.4809783
164. Faure A, Naldi A, Chaouiya C, Thieffry D. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*. Oxford University Press; 2006;22: e124–e131. doi:10.1093/bioinformatics/btl210
165. Naldi A, Monteiro PT, Christoph M, Models L, Kestler HA, Thieffry D, et al. Cooperative development of logical modelling standards and tools with CoLoMoTo. *Bioinformatics*. 2015; 1–7.
166. C. Chaouiya, A. Naldi DT. Logical Modelling of Gene Regulatory Networks with GINsim. *Methods in Molecular Biology*,. 1st ed. Springer; 2012. pp. 463–479. Available: [http://dx.doi.org/10.1007/978-1-61779-361-5\\_23](http://dx.doi.org/10.1007/978-1-61779-361-5_23)
167. Naldi A, Remy E, Thieffry D, Chaouiya C. Dynamically consistent reduction of logical regulatory graphs. *Theor Comput Sci*. 2011;412: 2207–2218. doi:10.1016/j.tcs.2010.10.021

168. Remy E, Rebouissou S, Chaouiya C, Zinovyev A, Radvanyi F, Calzone L. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. *Cancer Res.* 2015;75: 4042–4052. doi:10.1158/0008-5472.CAN-15-0602
169. Grieco L, Calzone L, Bernard-Pierrot I, Radvanyi F, Kahn-Perlès B, Thieffry D. Integrative modelling of the influence of MAPK network on cancer cell fate decision. *PLoS Comput Biol.* 2013;9: e1003286. doi:10.1371/journal.pcbi.1003286
170. Steinway SN, Zañudo JGT, Michel PJ, Feith DJ, Loughran TP, Albert R. Combinatorial interventions inhibit TGF $\beta$ -driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *npj Syst Biol Appl.* 2015;1: 1–12. doi:10.1038/npjsba.2015.14
171. Cohen DPA, Martignetti L, Robine S, Barillot E, Zinovyev A, Calzone L. Mathematical Modelling of Molecular Pathways Enabling Tumour Cell Invasion and Migration. *PLoS Comput Biol.* 2015;11. doi:10.1371/journal.pcbi.1004571
172. Flobak Å, Baudot A, Remy E, Thommesen L, Thieffry D, Kuiper M, et al. Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling. *PLoS Comput Biol.* 2015;11. doi:10.1371/journal.pcbi.1004426
173. Shin S-Y, Rath O, Zebisch A, Choo S-M, Kolch W, Cho K-H. Functional roles of multiple feedback loops in extracellular signal-regulated kinase and Wnt signaling pathways that regulate epithelial-mesenchymal transition. *Cancer Res.* 2010;70: 6715–6724. doi:10.1158/0008-5472.CAN-10-1377
174. Lu M, Kumar M, Levine H, Onuchic JN, Ben-jacob E. MicroRNA-based regulation of epithelial-hybrid mesenchymal fate determination. *Proc Am Thorac Soc.* 2013;110: 18144–18149. doi:10.1073/pnas.
175. Tian X-J, Zhang H, Xing J. Coupled reversible and irreversible bistable switches underlying TGF $\beta$ -induced epithelial to mesenchymal transition. *Biophys J. The Biophysical Society;* 2013;105: 1079–89. doi:10.1016/j.bpj.2013.07.011
176. Huang B, Jolly MK, Lu M, Tsarfaty I, Ben-Jacob E, Onuchic JN. Modeling the Transitions between Collective and Solitary Migration Phenotypes in Cancer Metastasis. *Sci Rep.* 2015;5: 17379. doi:10.1038/srep17379
177. Rateitschak K, Kaderali L, Wolkenhauer O, Jaster R. Autocrine TGF- $\beta$ /ZEB/microRNA-200 signal transduction drives epithelial-mesenchymal transition: Kinetic models predict minimal drug dose to inhibit metastasis. *Cell Signal.* 2016;28: 861–870. doi:10.1016/j.cellsig.2016.03.002

# Logical Modelling of the Regulation of Cell Adhesion properties in EMT





## Summary

E-cadherin mediated cell-cell adhesion and the focal adhesion are two critical features that change during EMT, and also characterize partial EMT phenotypes with collective migration properties (Hybrid phenotypes). The switching between Epithelial, Mesenchymal and Hybrid phenotypes is believed to play a key role in the process of metastasis. However, the microenvironment conditions that allow the control of the switching between phenotypes accounting for cancer mutational background are still unknown. The regulation of cell adhesion properties during EMT is complex, involving multiple microenvironment signals and signalling pathways. Mathematical modelling with the logical formalism has been demonstrated to be a useful tool to understand complex systems such as signalling in cancer. Based on an extensive literature search, we constructed a logical network model that encompasses a total of 8 microenvironment signals involved in the regulation of EMT (3 signals from cell-cell contacts, 4 from diffusive molecules, and 1 biophysical property from the ECM). We developed the model to account for the main regulatory signalling pathways represented by 42 intracellular components (nodes) and 134 regulatory interactions. In our model, we define the cell-cell adhesion and focal adhesion dynamics as model readouts, which were associated to the Epithelial, Mesenchymal and Hybrid phenotypes. Our model was able to recapitulate the stability of Epithelial and Mesenchymal phenotypes together with their reported markers. Further, the model recapitulated the observed activity changes of cell adhesion and molecular components from published experiments on Epithelial cell lines exposed to microenvironmental signals or containing mutations. With our model, we generated predictions suggesting that cell-cell contact signals that activate RPTP and FAT4 determine cell adhesion properties and control the switching between phenotypes. Further, we also predicted that SRC and FAT4 overexpression in carcinomas were two

critical perturbations that promotes the switching towards Mesenchymal and Hybrid phenotypes, respectively. Moreover, we also provided a mechanistic explanation for the effects of RPTP and FAT4 on the control of cell adhesion properties.

## **1. Introduction**

Epithelial-to-Mesenchymal transition (EMT) is a complex and reversible (MET) cellular program, where an Epithelial cell loses the adhesion to its neighbour cells (cell-cell adhesion) and transforms into a Mesenchymal cell with migration capacity [1–3]. This process is known to be hijacked by carcinomas, which is often associated with the acquisition of invasion capacity [4–6]. However, EMT in carcinomas is often incomplete (partial EMT), where the E-cadherin mediated cell-cell adhesion is maintained [2,7]. In particular cases, partial EMT in carcinomas results in the gain of the Mesenchymal migration capacity often called as Hybrid phenotypes [8–10]. Thus, carcinomas can migrate as single cells (e.g. Mesenchymal phenotype) or as a group of cells through the formation of Hybrid phenotypes [8,9,11]. Epithelial, Mesenchymal and Hybrid phenotypes can be distinguished by their characteristic E-cadherin mediated cell-cell adhesion in combination with the characteristic focal adhesion dynamics, two cell adhesion properties critical to ensure the mode of cancer migration [10,12–14]. For example, an intermediate strength of cell-cell adhesion is often found to be a feature of collective migration in carcinomas, whereas Epithelial cells have high cell-cell adhesion strength [10,12,15]. On the other hand, focal adhesion dynamics is an integrin mediated cell-matrix adhesion property that mediates the migration capacity along the extracellular matrix (ECM) of Mesenchymal cells [10,12,13]. Thus, we focused our work on these two cell adhesion properties as a proxy of Epithelial, Mesenchymal and Hybrid phenotypes.

The dynamics of the interconversion (switching) between Epithelial, Hybrid and Mesenchymal phenotypes are key steps that participate in the initial and final steps of metastasis of carcinomas [11,12,15–18]. The tumour microenvironment is considered as the main hypothesis that drives the switching between phenotypes in metastasis [11,16,17,19]. However, how and which microenvironment signals control the switching between phenotypes are still unknown [8,11,20]. This is mainly due to the complexity of the tumour microenvironment, which is composed by a variety of diffusible signals from the ECM and signals from neighbouring cells (cell-cell contact signals) in a dynamic equilibrium [8,16,19,21]. In vitro, several microenvironment signals such as ECM stiffness, inflammatory signals, hypoxia, growth factors and Delta-Notch cell-cell contact signals have been already postulated to play a key role in driving EMT in cancer cells [8,22–24]. On the other hand, microenvironment signals that counteract EMT are still unknown [19,25]. The regulation of EMT by the microenvironment is complex, involving multiple and intertwined signalling pathways with feedback loops [25–27]. The complexity further increases when mutations in cancer progression can alter the regulatory control [8,28]. Thus, innumerable hypotheses that consist of the combination of microenvironment signals with mutational effects can arise for the switching conditions between phenotypes. Accessing these conditions is a challenge suitable for systems biology frameworks, provides a methodology that facilitates the analysis and understanding of such complex systems and generate new testable predictions (hypothesis) [29]. Some kinetic and logical models of the regulation of EMT have been developed to address particular questions and to explain EMT properties [30–34]. Recently, a logical network model of EMT in hepatocarcinoma predicted inhibitory effects on intracellular signalling components capable of stabilizing partial EMT phenotypes by only considering E-cadherin (ON/OFF) as the

indicator of EMT [35]. However, the relation between multiple microenvironment signals and the acquisition of invasive adhesion properties (cell-cell adhesion and focal adhesion dynamics) on these phenotypes are yet to be explored. This motivated us to develop a network model to predict the microenvironment and mutational conditions for the switching between adhesion features that characterize Epithelial, Mesenchymal and Hybrid phenotypes. Because the kinetic laws and constants are unknown for most signalling processes, the logical modelling framework is an alternative systems biology approach since it does not require knowledge of the kinetics [36]. This framework has been quite successful in recapitulating experimental observations and generating robust and testable predictions in cancer research [31,32,37–39]. Thus, we chose to develop a literature-based logical network model for the regulation of E-cadherin mediated cell-cell adhesion and focal adhesion dynamics. Here, we present a model for the regulation of cell adhesion properties that accounts for the main microenvironment signals and signalling pathways that drive EMT. This model was validated by recapitulating Epithelial and Mesenchymal markers, and observations from experiments on Epithelial cell lines that included the effects of several microenvironment signals or mutations. In this chapter, we focus on the analysis of the role of cell-cell contact signals on the regulation of cell adhesion properties during EMT. Here, we explore the potential of cell-cell contact activation of Receptor Protein Tyrosine Phosphatases (RPTP) and the membrane receptor FAT4 to control EMT [40,41]. These receptors were chosen based on their reported potential to inhibit RTK (RPTP) and Wnt signalling through Hippo signalling activation (FAT4), signalling pathways associated with EMT progression [40,42]. Our model analysis resulted in predictions for the capacity of RPTP and FAT4 activation by cell-cell contacts to control cell adhesion in the presence of all possible combinations of EMT driving signals. In this work, we

also identified critical alterations in the network that favoured either Mesenchymal or Hybrid adhesion phenotypes. Among them, two correlate with tumour expression degree in comparison with normal tissues, and their roles in the switching of phenotypes was explored. Moreover, we provide here a mechanistic explanation based on our model for the effects of RPTP and FAT4 signals on cell adhesion properties during EMT.

## **2. Methods**

### **2.1. Software tools**

#### **2.1.1. GINsim**

GINsim is a free software tool developed in JAVA for the modelling and simulation of regulatory networks ([www.ginsim.org](http://www.ginsim.org)) [47]. This tool is the result of a collaborative development coordinated by C. Chaouiya (Instituto Gulbenkian de Ciência, Oeiras) and D. Thieffry (Ecole Normale Supérieure, Paris). We used the currently stable version 2.4 for the setup and analysis of the model presented in this thesis. We also used a version under development with new analytic tools for attractors search, implemented by Rui Henriques at INESC-ID and supervised by Pedro Monteiro and C. Chaouiya.

#### **2.1.2. R and Rstudio**

For numerical manipulations and data plot, we used R version 3.3.2, a free software environment for statistical computing and graphics ([www.r-project.org](http://www.r-project.org)) initially created by Ross Ihaka and Robert Gentleman in 1993. R code was written and run under the Rstudio® integrative console environment, version 1.0.136.

### **2.1.3. Python**

For automatic and systematic analysis of results, we wrote several scripts using the python version 2.7 ([www.python.org](http://www.python.org)), an object oriented high-level programming language, originally created by Guido van Rossum in 1991.

### **2.1.4. Scripts**

R and python scripts used in the work presented in this chapter are available in GitHub public repository (<https://github.com/rjpais>) under the project cell adhesion network model. We include the running instructions inside each file.

## **2.2. Model construction**

### **2.2.1. Network construction**

To construct and visualize the network, we used GINsim software [47]. With this tool, we built the interaction map (network) which contains the regulators (nodes) connected through arcs, representing activations or inhibitions (interactions). We based on regulators and interactions from signalling pathways and regulatory processes involved in cell-cell adhesion and focal adhesion to construct the network. To find these regulators and interactions, we started from records in KEGG PATHWAY database [48]. Initially, we explored the KEGG reference maps of adherens junction, focal adhesion and several signalling pathways that participate in cell adhesion regulation (Table 3). To complement this source, we used the Atlas of cancer signalling database (ACSN) developed at the Institute Curie (<https://acsn.curie.fr>). In particular, we explored the ACSN map for EMT and cell motility that contains modules for the regulation of cell-cell and cell-matrix adhesion (Focal adhesion) [49]. We used the NaviCell web platform to navigate on the ACSN and find additional interactions and regulators [50]. We also brought-in the regulators and interactions of the EMT network developed for

hepatocytes by Steinway et al. [31]. Next, we searched the literature for experimental work, preferably on Epithelial cell lines, in order to provide support for each interaction in the network. Additional interactions not found in databases were obtained from references found in the literature.

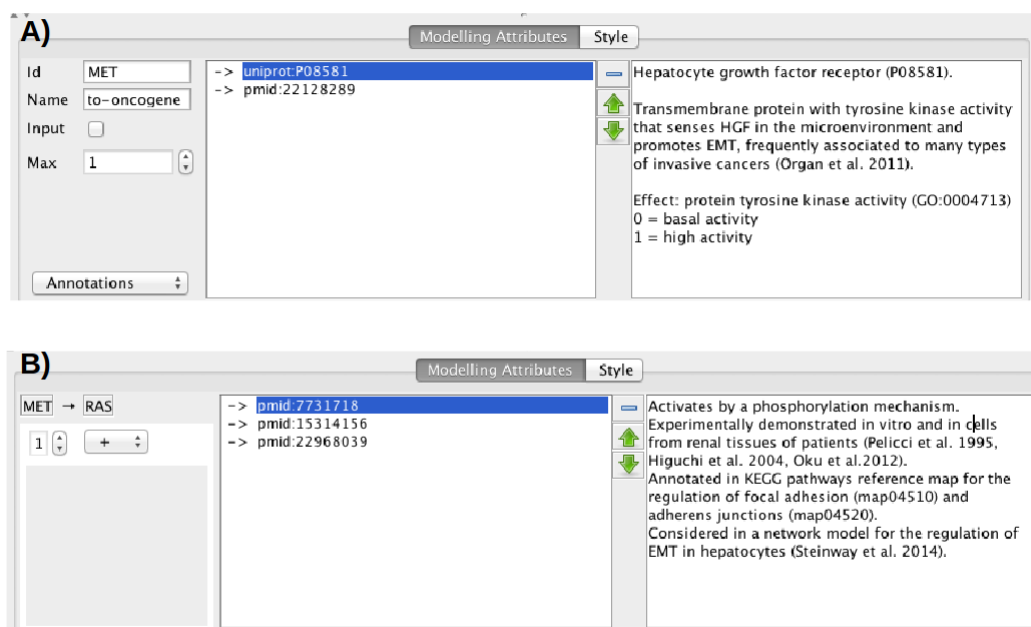
**Table 3 .KEGG reference maps used for network construction.** Maps were taken from KEGG PATHWAY database available at [www.genome.jp/kegg/pathway.html](http://www.genome.jp/kegg/pathway.html). Maps indicated with \*\*\* were used for the main basis of the regulatory network, \*\* additional pathways included, \* maps inspected to find additional crosstalk points.

Process/signalling pathways	Reference maps for humans	Usage	Last update
Adherens junction	hsa04520	***	2013
Focal adhesion	hsa04510	***	2015
Notch signalling	hsa04330	**	2017
Hippo signalling	hsa04390	**	2017
HIF1 signalling	hsa04066	*	2017
JAK-STAT signalling	hsa04630	*	2015
TGF signalling	hsa04350	*	2015
Wnt Signalling	hsa04310	*	2017
AKT signalling	hsa04151	*	2017
MAPK signalling	hsa04010	*	2017
RAP1 signalling	hsa04015	*	2015
ERBB signalling	hsa04012	*	2015
NF-kappa B signalling	hsa04064	*	2017
RAS signalling	hsa04014	*	2017

### 2.2.2. Network annotations

For each node of the network, we annotated the name of the correspondent molecular component, type of molecule, cellular localization, role in cell adhesion and molecular activity. We used the UniProt ([www.uniprot.org](http://www.uniprot.org)) and the NCBI gene database ([www.ncbi.nlm.nih.gov/gene](http://www.ncbi.nlm.nih.gov/gene)) as sources for the annotation. For each molecular component, we check on these databases the tissue specificity, existence in humans, and the Gene Ontology (GO) term that would best describe

the molecular activity. When needed, specific reviews on each component were used to get further biological information. For interactions, we collected information about the biological mechanisms from the supporting reference articles. We also searched on the supporting references for information regarding cooperative relations with other model components. We defined a controlled vocabulary to simplify the description of the biological mechanisms of the interactions (Table 4). Because these mechanisms have distinct time scales, we also annotated the correspondent time scale of the interaction based on the slowest biological mechanism identified. We stored all annotations in GINsim annotation box, where we included the links to references using PMID, UniProt and NCBI identifiers (Figure 16) [47]. Here, we also included additional comments about the presence of the interaction in signalling databases or published model.



**Figure 16. Examples of annotations in GINsim annotation box. (A)** Annotation for the node MET. **(B)** Annotation of the interaction between MET and RAS.

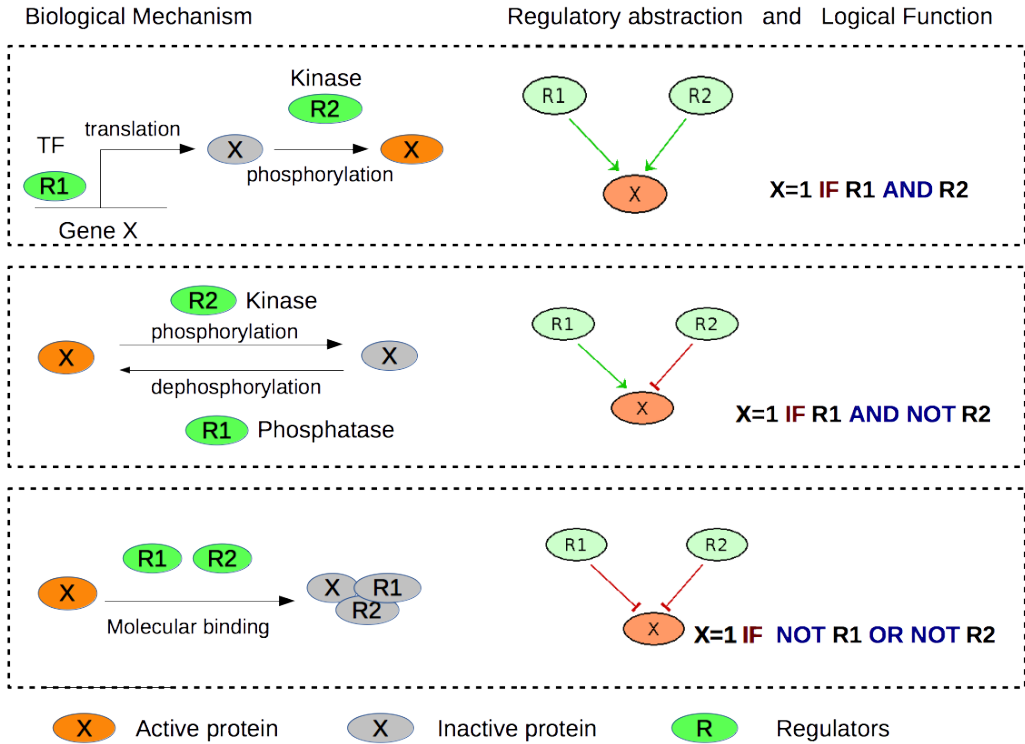


**Table 4. Controlled vocabulary used for the description of biological mechanisms and their correspondent time scales.**

Biological mechanism (interactions)	Time scale	References
Transcription	SLOW (minutes)	[51,52]
Translation	SLOW (minutes)	[51,52]
Oxidation	SLOW (minutes)	[53,54]
Phosphorylation	FAST (seconds)	[55,56]
Dephosphorylation	FAST (seconds)	[55,56]
Molecular binding	FAST (seconds)	[51,57]
Transport	FAST (seconds)	[51,58]

### 2.2.3. Logical model construction

For the exploration of the regulatory capabilities of the constructed network, we setup a logical model in GINsim based on the formalism initially proposed by René Thomas (see sections 3.1 and 3.3 of chapter 1 for formalism details) [47,59]. First, we associated a Boolean (0 or 1) or multi-valued (0, 1, 2) variable to each node of the network. For the nodes with no incoming interactions, we set them as model inputs. For each of the remaining nodes, we set a logical function for the value 1 and, if multi-valued, we set another for value 2. No function was required for value 0 in GINsim because this value is used by default when function evaluations are false. Logical functions were encoded by connecting regulators (R) of a target node (X) through the logical operators OR (denoted in GINsim by |), AND (&) and NOT (!). Logical functions were chosen to abstract the biological mechanisms of activation and inhibition of model components, annotated in the network (see illustrated examples in Figure 17). Here, inhibitory roles were treated as the negation of regulator role (NOT R). Finally, we checked the logical functions for consistency using a tool implemented in GINsim that computes the interaction functionality.



**Figure 17. Examples of biological mechanisms and their correspondent logical abstraction.** Target nodes are identified by X and regulators by R1 and R2. Active regulators are depicted in green, active target in orange and inactive target in grey. The logical abstraction of each mechanism (left) is composed by the regulatory graph and the logical functions (right). In the first mechanism (top), a transcription factor (R1) binds to the gene promoter and activates the translation of the protein X, which requires being phosphorylated by a kinase (R2) to activate its function. The second example (middle) shows the modulation of an protein activity by the phosphorylation/dephosphorylation degree of the active site, mediated by kinases (R2) and phosphatases (R1). The last example (bottom) shows the activation of a protein by the formation of an active complex with other two molecules (R1 and R2).

#### 2.2.4. Definition of updating priority rules

To generate the model dynamics accounting for mechanisms that occur in two timescales (seconds and minutes), we defined two priority classes for the update of variables (FAST and SLOW) [60]. We defined FAST as a set of nodes that update first and SLOW the set of nodes that update last. We defined FAST and

SLOW sets for two updating situations, when a node value increases (+) and when it decreases (-). We set the priority rules in GINsim dialog box (Tools / Run simulation / Priority Class Selection / Configure) [47]. Here, whenever there is a call to update variables of nodes, the model will first update the variables of the nodes that belong to the list of first priority (FAST) if any. Variables will then iteratively update asynchronously, generating successive co-current states. In each update call, logical functions are re-evaluated to match each of the resulting states. Only when no variable of the first priority nodes is required to update, the variables of the nodes of the second priority (SLOW) are updated. This iterative update stops when logical functions are no longer updatable, reaching one or more stable states. We selected the nodes for FAST and SLOW sets based on the identification of slower mechanisms (e.g. transcription) as limiting steps on the annotated regulatory interactions (section 2.2.2.).

#### **2.2.5. Model documentation**

To facilitate the access to the model documentation and references, we generated an HTML file (Model\_documentation.html) that contains a brief description of the model, the components (nodes), logical functions and annotations. In this file, we add the links to the abstracts of each reference through the PubMed or UniProt ID numbers. To generate this file, we used the export tool functionality implemented in GINsim (File/export/Documentation). The file was made available in the GitHub public repository (<https://github.com/rjpais>) under the project cell adhesion model. After the download of the file, run the file to visualize in your web browser. In alternative, <http://htmlpreview.github.io/MD>).

## **2.3. Model static analysis**

### **2.3.1. Computing model stable states**

Regardless of their reachability properties, we analytically determine the model stable states using an algorithm based on multi-valued decision diagrams, developed and implemented in GINsim by Naldi et al. [61]. We computed the stable states using python scripts adapted from the one developed by Naldi and provided by the lab of C. Chaouiya. First, we adapted the python script to compute the stable states of the unperturbed model (GetModelSS.py). Next, we adapted the python script (GetModelSS\_SP.py) to systematically compute the stable states for each perturbation on a single model component (node), where the value of the node (0,1, or 2 if multi-valued) was fixed to be constant (single perturbation). Finally, we systematically computed the stable states of all combinations of single perturbations with the python script from Naldi (GetModelSS\_SDP.py). The computed stable states for unperturbed, single and double perturbations were saved in 3 separated text files to be further processed in the subsequent static analysis. To generate a separate file with the stable states for each perturbation, we wrote a script that extracts the stable states of a selected perturbation from the files that contain single or double perturbations (SelectPerturb.py).

### **2.3.2. Identification of the stable states of adhesion phenotypes**

For the analysis of the activity patterns of each adhesion phenotype, we wrote a python script (InternalSS.py) to identify distinct stable states of the model excluding the model input values (internal stable states) for each model output combination (phenotype). To perform the identification, we run the script to parse the previously generated files with the stable states of the unperturbed model

(section 2.3.1.) and perturbations (section 2.3.2). The script was written to remove the input nodes from the stable states and systematically identify distinct stable states (only with the internal nodes), discarding repetitions. Next, the stable states were ordered by output combinations and printed to an output file. With this script, we also automatically checked if the value of each node in the stable states was conserved (if has the same activity level) for each adhesion phenotype. Then, a stable state showing the nodes that had conserved activity and the ones that had different activity degrees was computed and printed to the output file.

### **2.3.3. Single and double perturbation analysis**

To identify the effects of single perturbations, we developed a python script (SP\_analysis.py) to parse the file with the stable states of these perturbations (section 2.3.1) and identify the perturbations that result in the gain or the loss of an adhesion phenotype. With this script, we systematically identify the output combinations (adhesion phenotypes) within the stable states for each perturbation. Next, the presence or absence of each output combination was systematically compared with the ones computed for the unperturbed model (reference). This allowed us to identify automatically a gain or a loss of an adhesion phenotype. For the analysis of double perturbations, we adapted the previous script to identify only the double perturbations that result in a synergistic effect (not obtained in the single perturbations). We used the modified script (SP\_analysis.py) to parse the file with the single and double perturbations and systematically compare the gains and losses of adhesion phenotypes between double and single perturbations. We set filter conditions in this script to select the double perturbations that had a gain/loss of phenotype not verified in the case of the two single perturbations. To show the results, the output file was printed in the

form of lists containing all perturbation names that were identified to cause a gain/loss of adhesion phenotypes, organized by phenotype and effects (gain/loss).

#### **2.3.4. Identification of input conditions for the adhesion phenotypes**

To systematically analyse the compatibility of adhesion phenotypes with microenvironments, we wrote a python script (MicPhenComb.py) to identify the combinations of model input values (microenvironments) for each combination of output values (adhesion phenotype) on the model stable states. This script was designed to parse a file with the stable states and returns a file with a list of combinations of input values (without repetitions) for each adhesion phenotype. With this script, the input combinations were listed for each combination of values of two selected model inputs under analysis (RPTPL and FAT4L in this work). We also compute the percentage of the remaining input combinations for each combination of RPTPL and FAT4L values for each adhesion phenotype. We used the previously produced files (section 2.3.1.) containing the stable states of the unperturbed model and selected perturbations as input files to run the script.

### **2.4. Analysis of model dynamics**

#### **2.4.1 Simulation of biological scenarios**

For the simulation of biological scenarios, we first set the model input values in GINsim toolbox (Tools / Run simulation) according to the microenvironment conditions described in experiments (model validation) or to mimic a physiological scenario (model analysis). When needed, we set perturbations on model components in the dialogue box (Tools / Run simulation/ Select a perturbation / Configure). These perturbations were set to mimic reported knockouts or

overexpression of genes that produce molecular components of the network. For knockout effects, we fix the node value to 0. For an overexpression that results in an ectopic activation, we chose to fix the node value to 1 or 2 if multi-valued. In the case of an overexpression that is dependent on a subsequent activation by a regulator and is subject to an inhibition by another regulator, we choose to knockout the inhibitory effect by fixing the value of the regulator effect to 0. This is a more realistic perturbation for the cases where the overexpression of a protein leads to the insensitivity to its inhibitor (e.g. competitive inhibition), where the increase in concentration reduces the effect of the inhibitor, but yet the protein still requires to be activated by a post-translation modification (e.g. phosphorylation) or other molecular processes such as molecular binding. To analyse the reachability of adhesion phenotypes, we set the simulations to start from all possible combinations of initial states (from all state space). To analyse the dynamics of transitions between phenotypes, we set the values of the nodes in the initial state of simulations to correspond the stable state of a particular phenotype (e.g. Epithelial, Mesenchymal or Hybrid) compatible with the conditions before the perturbation. Simulations were then performed with two complementary methods described in the sections 2.4.2 (transitions between phenotypes) and 2.4.3 (reachability of phenotypes).

#### **2.4.2. Simulation of the state transition graph**

To obtain the stable states that can be reached starting from a particular biological scenario (section 2.4.1), we computed the state transition graph (STG) using the “run simulation” tool implemented in GINsim [47,62]. In these graphs, the nodes are the vectors encompassing the activity values of model components (states), while the arcs connect successive states, called “state transitions” (see section 3.4 of chapter 1 for more details). In the STG, a stable state is found

when a state with no successors is reached. This method enables the exploration of all trajectories according to an updating scheme, represented through the STG [62]. We have run simulations under asynchronous updating policy, according to rules of priority that account for timescale constraints of the regulatory mechanisms included in the network (section 2.2.4.) [60,62]. We believe that with the priority rules, we explore only the biologically plausible “state transitions” which results in a reduced and more realistic STG.

### **2.4.3. Probability of reaching stable states**

To estimate the probabilities of reaching stable states or adhesion phenotypes starting from a particular biological scenario (section 2.4.1), we used the tool “Analyze attractors” implemented in GINsim. We chose the MonteCarlo algorithm to perform a randomized search for stable states, exploring a subset of all possible “state transitions”. This algorithm was chosen over others that perform a more exhaustive and detailed search for several types of attractors (AVATAR and FireFront). Since we were only interested in the reachability of stable attractors, the MonteCarlo algorithm was chosen because is faster and as effective as AVATAR or FireFront for the estimation of the probabilities of reaching stable states. To estimate the probabilities of the reaching the stable states with MonteCarlo algorithm, we have run simulations with a total of 10000 iterations that explored a maximal number of 10000 “state transitions” randomly visited in each iteration (depth of trajectories).



## 2.5. Model validation data and comparisons

### 2.5.1. Epithelial and mesenchymal markers

To validate the model stable states of adhesion phenotypes, we used Epithelial and Mesenchymal markers commonly used as markers of EMT, because these cell types have distinct cell adhesion properties [63–65] (Table 4). The used markers consist of proteins or RNA that are found to be differentially expressed in Epithelial and Mesenchymal cell types in specific cellular localization.

**Table 4. EMT biomarkers used for model validation.**

Molecule / marker	Cellular Location	Expression in Epithelial cells	Expression in Mesenchymal cells	References
E-cadherin	Membrane	Present	Absent	[63–65]
miR200	Cytoplasm	Present	Absent	[63,64]
E-cadherin/ $\beta$ -catenin complex	Membrane	Present	Absent	[64,65]
E-cadherin/p120	Membrane	Present	Absent	[64,65]
Snail 1 and snail 2	Nuclear	Absent	Present	[63–65]
Zeb 1 and zeb 2	Nuclear	Absent	Present	[63–65]
B-catenin	Nuclear	Absent	Present	[63–65]
LEF1	Nuclear	Absent	Present	[63,64]

### 2.5.2. Experiments on Epithelial cell lines

To test the model, the scientific literature provides reports with experiments on mammalian Epithelial cell lines exposed to microenvironment signals and/or containing mutational perturbations (e.g. knockout or overexpression of genes). We selected 28 published reports, where we collected a total of 163 observations, 26 on cell-cell adhesion, 12 on focal adhesion and 125 on the molecular activity of molecular components used in the model. In these selected

experiments, we did not include published work used as references for the network construction. We grouped these selected experiments into 13 sets (Table 5) that consider the same conditions (microenvironments and/or mutations).

**Table 5. Selected experiments on mammalian epithelial cell lines used for model validation, grouped in perturbation type.** The number of total observations (N) collected for each perturbation type.

<b>Perturbation</b>	<b>Experiments description</b>	<b>N</b>	<b>Ref.</b>
<b>IL6+</b>	Exogenous addition of IL6 to normal mammary Epithelial cell lines.	5	[74]
<b>EGF+</b>	Exogenous addition of epidermal growth factor to several human Epithelial cell lines (cancer and normal).	18	[75-78]
<b>ECM+</b>	Exogenous addition of collagen type I or fibronectin to several human and mouse Epithelial cell lines (cancer and normal).	10	[78,79]
<b>EGF+ ECM+</b>	Exogenous addition of epidermal growth factor in combination with collagen type I or fibronectin to human breast, skin and squamous Epithelial cell lines (cancer and normal).	5	[78]
<b>HGF+</b>	Exogenous addition of hepatocyte growth factor to Epithelial cell lines from human, canine and mouse origins.	35	[33,76,80-82]
<b>WNT+</b>	Exogenous addition of Wnt protein to kidney canine and colon cancer Epithelial cell lines.	8	[33,83]
<b>TGFb+</b>	Exogenous addition of transforming growth factor beta to several human and mouse Epithelial cell lines.	30	[84-87]
<b>EGF+ ROS+</b>	Exogenous additions of epidermal growth factor with ROS intracellular generation in mammary Epithelial cell lines.	10	[88]
<b>P120 KO</b>	Knockout of p120 using siRNA in hamster ovary cell lines.	8	[66,67]
<b>B-Cat_mem KO</b>	Transformation of mouse embryonic and canine kidney Epithelial cell lines with a mutated form of E-cadherin incapable of binding to $\beta$ -catenin.	5	[33,68]
<b>CK1 E1</b>	Overexpression of CK1 on human endothelial and canine kidney cell lines.	7	[69]
<b>CD44 E1</b>	Overexpression of CD44 in colon tumour cancer cell lines.	5	[70]
<b>SRC E1</b>	Overexpression of SRC in human breast and colon cancer cell lines.	17	[71-73]

For cell adhesion properties we collected observations from cell adhesion assays or imaging techniques on these reports. For the observed activity of molecular

components, we used the data from quantitative assays such as qPCR, ELISA and western blotting of the active forms (e.g. phosphorylated protein). We analysed and interpreted the collected observations in a qualitative manner by comparing the changes in activity with the experiment control (Epithelial cell lines not exposed to external or internal perturbations). Finally, we discretise the activity of cell adhesion properties and molecular components in 3 qualitative levels: increased (+), decreased (-), or not changed (~).

### **2.5.3. Matching simulations to experimental observations**

To test the compatibility of the results from the simulation of experiments (stable states) with the experimental observations, we constructed two matrices with the same dimensions (number of nodes x number of experiments), one for the experimental observations and another with the stable states from simulations. Because the experimental observations contained 3 qualitative levels of activity (+, 0, -), we manipulated the matrix of simulations to compute a matrix with the variations in sign (+, - or 0). Here, we subtracted the values of the final state (stable state) to the ones of the initial state used for simulations. Next, we compared the two matrices and computed a matrix that identified the differences in sign. All manipulations of the matrices were performed using R studio.

### **2.5.4. Model comparison with a published EMT network model**

To further compare the stable states of Epithelial and Mesenchymal adhesion phenotypes of our model with the reported Epithelial and Mesenchymal activity status, we used the published EMT network model for hepatocytes developed by Steinway et al. [31]. First, we setup the model of Steinway et al. in GINsim using the model components, interactions and logical rules provided in [31]. Then, we computed the Epithelial and Mesenchymal stable states for this model in GINsim

as it is described in sections 2.3.1. and 2.3.3. for the unperturbed model. To compare the stable states of both models, we ordered the common nodes in GINsim and removed the nodes not common from the files in the stable states.

### **2.5.5. Protein expression in carcinomas.**

To identify relevant predictions made by our model, we observed the expression data of carcinomas from breast, prostate, colorectal and lung tissues. We used the human protein atlas database to find this data ([www.proteinatlas.org](http://www.proteinatlas.org)). This database provides frequencies of expression on a qualitative scale used to compare with our qualitative predictions. We used the data provided in this database which was experimentally determined by antibody staining of carcinoma tissue samples, already statistical analysed in comparison with normal tissues [89]. Therefore, for each model component that is a protein, we searched on this database for the relative expression levels in carcinoma tissues [90]. We used this data to identify overexpressions and knockouts of model components in carcinomas that correlated with our predictions generated for single and double perturbations.

### **2.5.6. Data plots**

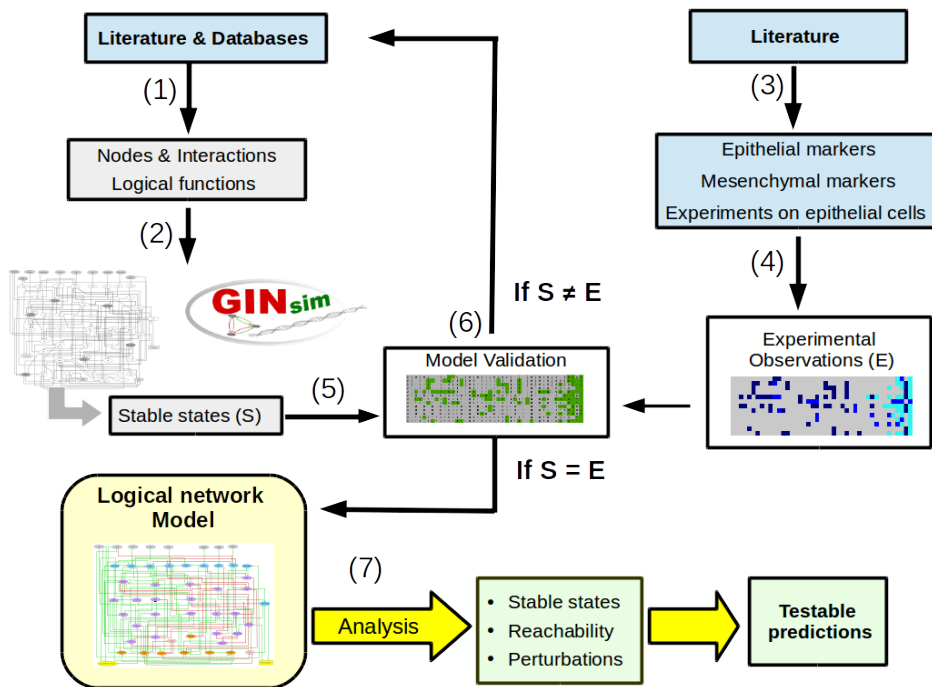
For the visualization of the model stable states, observations of experiments on epithelial cell lines and perturbation effects, we adapted the `heatmap2()` function in R (see files `Heatmap_validation.R`, `Heatmap_experiments.R`, `InternalSS.R`, and `perturbations_effects.R`). To show the results of the microenvironments counts, we used grouped bar plots defined in R with the `barplot()` function (see file `Barplot_Phenotypes.R`). For showing the probabilities of reaching stable states we used pie charts defined in R with the `pie()` function in R (see file `PieChart_Phenotypes.R`).

### 3. Results

#### 3.1. Modelling workflow

For our purpose, we developed and analysed a logical model following the workflow presented in Figure 18. We first constructed a regulatory network (step 1) based on the collected biological information from literature and databases (details in sections 2.2.1 and 2.2.2). Then, we defined a logical model (step 2) for the regulatory network with rules (logical functions) that obey the biochemical principals of each biological process modelled in the network. For this, we used the biological information annotated on the nodes and interactions (details in sections 2.2.3 and 2.2.4) together with reports that supported our choices. Independently from the literature data used for model construction, we collected qualitative observations from literature to test and validate the model (step 3). For this, we used reported markers for Epithelial and Mesenchymal cells (section 2.5.1), and reports containing experiments on Epithelial cell lines exposed to exogenous signals or with mutational alterations (section 2.5.2). The collected observations from experiments were then treated in a qualitative manner (step 4) such as we could compare them with the model stable states (see section 2.5.3). Next, we computed model stable states and simulated the experiments on Epithelial cell lines to compare with the experimental observations (step 5). We considered the model validated when all results were in agreement with the experimental observations (step 6). Here, no contradictory results were allowed in terms of qualitative changes in activity (e.g. reduction or increase). When we found such contradictions, we returned to the literature, revised the model interactions, functions and searched for missing interactions (repeat steps 1, 2, 5 and 6). Therefore, we started from a reduced network which evolved into a larger network model capable of explaining all experimental observations, and still

keeping the logical rules constrained by the supporting biological data. During the time of Ph.D., several model versions were produced that were not able to recapitulate all collected experimental observations. Some versions were able to recapitulate all data but did not account for regulatory interactions considered relevant for EMT and cell adhesion newly found during literature searches. Thus, we decided to discard these versions from this thesis assuming that the most complete and refined is more relevant to be analysed.



**Figure 18. Workflow for the logical network model construction, validation and analysis.** The numbers in parenthesis indicate distinct steps of the workflow that were followed.

After successive model refinements, we reached a final model version (presented in section 3.2) that recapitulated all experimental observations (sections 3.3 to

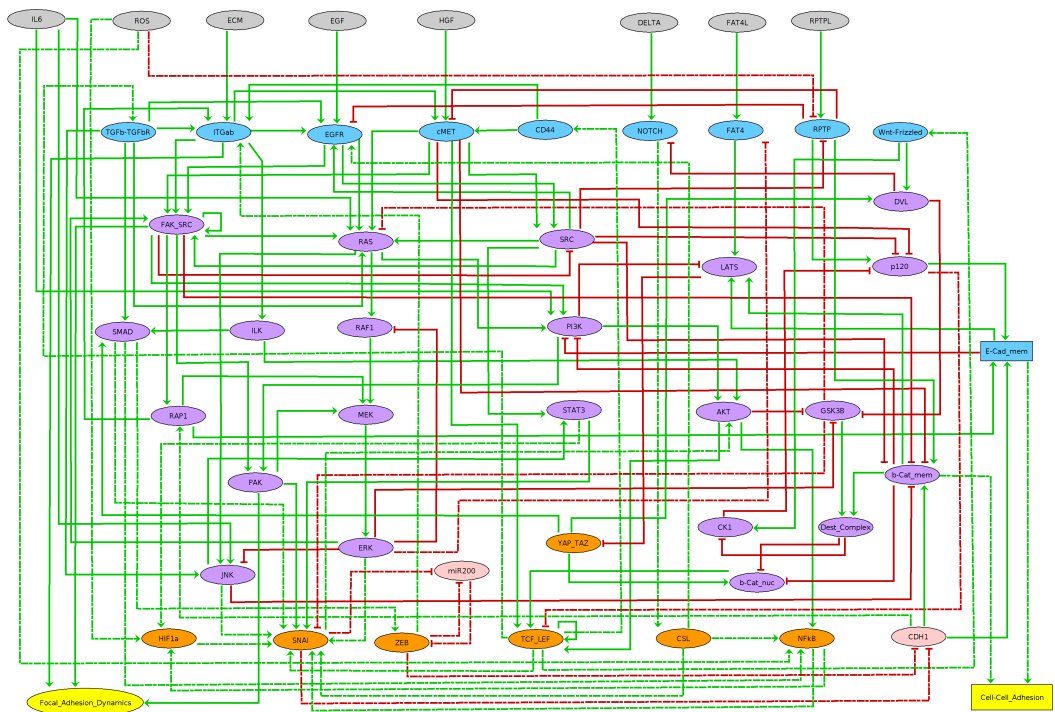
3.5). For this final version, we analysed the model dynamics in terms of its stable states to generate testable predictions of conditions that control cell adhesion properties during EMT (step 7). Therefore, we conducted a model analysis to establish the relationship between model inputs (microenvironment) and outputs (phenotypes), reachability properties of model outputs, and the effects of perturbations on model components (sections 3.6 to 3.10).

### 3.2. Logical network model of cell adhesion properties

The constructed network of the regulation of cell adhesion properties is presented in Figure 19 (Table S1 for nodes and Table S2 for interactions). For more details on the nodes, interactions, and respective references, see the `model_documentation.xhtml` file (see section 2.2.5 to get file). In our network, we considered the following signalling pathways and regulatory modules, where we represent them by key regulatory components:

- TGF $\beta$  signalling (TGFb-TGFbR  $\rightarrow$  SMAD);
- Integrin signalling (ILK  $\leftarrow$  ITGab  $\rightarrow$  FAK-SRC  $\rightarrow$  RAP1  $\rightarrow$  JNK);
- Wnt signalling (Wnt-Frizzled  $\rightarrow$  DVL  $\rightarrow$  GSK3B  $\rightarrow$  Dest\_Complex  $\rightarrow$  b-Cat\_nuc  $\rightarrow$  TCF/LEF);
- AKT signalling (PI3K  $\rightarrow$  AKT  $\rightarrow$  NFkB);
- MAPK signalling (RAS  $\rightarrow$  RAF1  $\rightarrow$  MEK  $\rightarrow$  ERK)
- HIF1 signalling (ROS  $\rightarrow$  HIF1a  $\leftarrow$  STAT3  $\leftarrow$  IL6/SRC);
- Notch signalling (DELTA  $\rightarrow$  NOTCH  $\rightarrow$  CSL);
- Hippo signalling (FAT4L  $\rightarrow$  FAT4  $\rightarrow$  LATS  $\rightarrow$  YAP\_TAZ);
- E-cadherin transcription regulators (SNAI, ZEB, miR20, CDH1);
- Phospho-regulation of E-cadherin/ $\beta$ -catenin complex (RPTP, MET, SRC, b-Cat\_mem, p120, E-cadherin\_mem).

In the regulatory network (Figure 19), we included several reported crosstalk interactions between the above mentioned signalling pathways and regulatory modules. Among them, we included: The Notch signalling crosstalk with the HIF1, TGF $\beta$ , MAPK, and AKT signalling pathways [91]; The Integrin and growth factor signalling crosstalk, including MAPK and AKT signalling [92–94]; and the Hippo signalling crosstalk with Wnt signalling [95,96]. For more details on the reasons to include the key regulatory interactions in the network and their underlying assumptions see the supplemental text on end of this thesis chapter (sections 5.1.1 to 5.1.3).



**Figure 19. Regulatory network of cell-cell adhesion and focal adhesion dynamics.** Microenvironment signals are depicted in grey, membrane protein/receptors in blue, signalling components in purple, transcription factors in orange, genes and RNA in pink, cell adhesion properties in yellow; elliptic nodes denote Boolean variables and rectangular nodes denote multi-valued variables; red arcs are inhibitions and green arcs activations. Interactions are classified into two distinct time scales: fast (e.g. phosphorylation) and slow (e.g. transcription and translation). Fast interactions are denoted by plain arcs and slow interactions by dashed arcs.



As model readouts, we defined 2 output nodes, one for the E-cadherin mediated cell-cell adhesion strength (Cell-Cell\_Adhesion), and another for the integrin mediated focal adhesion dynamics (Focal\_Adhesion\_Dynamics). For Cell-Cell\_Adhesion, we associate it to multi-valued variables taking the values 0, 1 or 2. These values mean basal (0), intermediate (1) and high (2) degrees of cell-cell adhesion strength to distinguish cell-cell adhesion degrees of Epithelial, Mesenchymal and partial EMT phenotypes [10].

For the Focal\_Adhesion\_Dynamics node, we chose to treat it as a Boolean variable for simplicity. In this case, the value 1 represents a high formation of focal adhesions with subsequent high turnover that represents the adhesion status of a migrating cell [13,97]. On the other hand, the value 0 means a basal focal adhesion formation or a basal turnover, which represents a non-migrating adhesion status. We considered a total of 8 microenvironment signals as model inputs (IL6, ROS, ECM, EGF, HGF, DELTA, FAT4L, and RPTPL). These input nodes were associated to Boolean variables defining basal (value 0) or high (value 1) degree, whereas value 0 have no effect and 1 have an effect on its targets (activation or inhibition). For the node that corresponds to E-cadherin in the membrane (E-Cad\_mem), we chose to associate it to a multi-valued variable to account for the basal (value 0), intermediate (1) and high (2) degrees of concentration [98–100]. For the remaining 40 nodes, we associate them to Boolean variables, where the value 0 and 1 means basal (no effect) and high degrees of activity (activation or inhibitory effect), respectively. For each node, the logical functions that define the evolution of the variables of each node (except inputs) are presented in the Table S3. The rational and support for the choices of logical rules are described in section 5.1.4 in the supplementary text of this chapter. In the supplemental text (section 5.1.5), we also provided a

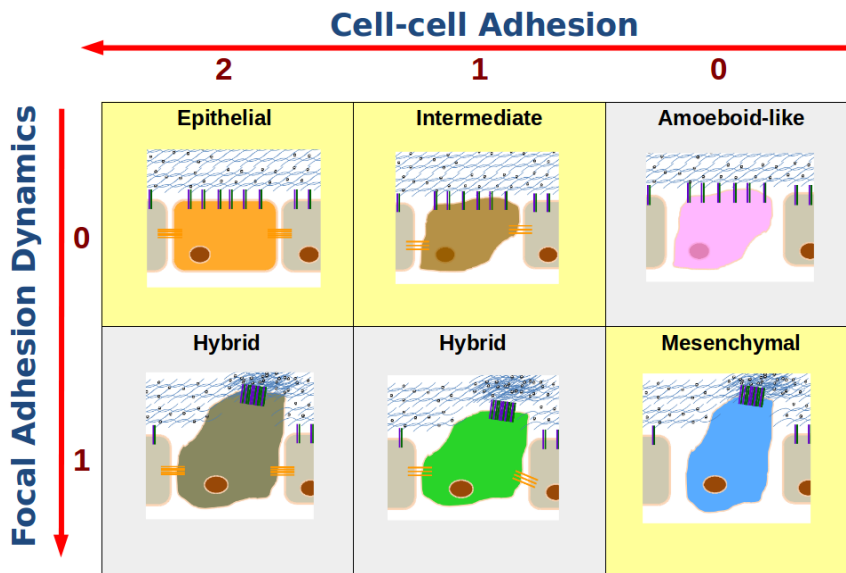
comparison of our model in terms of its similarity with a published EMT network model.

### **3.3. Recapitulation of model phenotypes**

To test the model capacity to recapitulate phenotypes, we check which combinations of cell adhesion properties were present in the computed model stable states (section 2.3.1), and which of those are characteristic of naturally observed phenotypes. In the model stable states, we obtained 3 out of 6 possible combinations of adhesion properties that correspond to distinct phenotypes (Figure 20). One was the combination of high cell-cell adhesion ( $C = 2$ ) with basal focal adhesion dynamics ( $F = 0$ ), which is characteristic of Epithelial cells [13,98,101]. Another combination was the basal cell-cell adhesion ( $C = 0$ ) with high focal adhesion dynamics, two properties characteristic of Mesenchymal cells [10,13,101]. Both Epithelial and Mesenchymal phenotypes are expected to be recovered in the stable states since they are naturally observed during EMT in several biological systems such as wound healing, development, and cell culture [2,102,103].

We also obtained stable states with the combination of intermediate cell-cell adhesion strength ( $C = 1$ ) and basal focal adhesion dynamics ( $F = 0$ ). We considered this adhesion combination to be compatible with an Intermediate phenotype that results from partial EMT. These adhesion properties are observed during collective migration of border cells during oogenesis, supporting this result [14]. This further indicates that we included in our model the microenvironment conditions that stabilize the Intermediate phenotype. Interestingly, the model stable states did not recover adhesion properties characteristic of Hybrid and Amoeboid-like phenotypes. This result is compatible with the idea that these phenotypes are not stable or transiently formed in normal cells, and their stability

is only reached during cancer progression [104,105]. Thus, we considered this result to be positive for supporting the model validation. Although our model contains an output combination that is compatible with the Amoeboid adhesion, our model is not suitable for this phenotype. The reasons are based on: **1)** Cell adhesion features are not suitable to correctly describe the invasive properties of this amoeboid-like migration [10,11,106]; **2)** Their relevance for metastasis is associated to other types of cancers such as leukaemia [11,15]; **3)** The Mesenchymal to Amoeboid transitions is dependent on the concentration of Rac1/RhoA activation, which is a regulatory mechanism not accounted in our model [105,106].



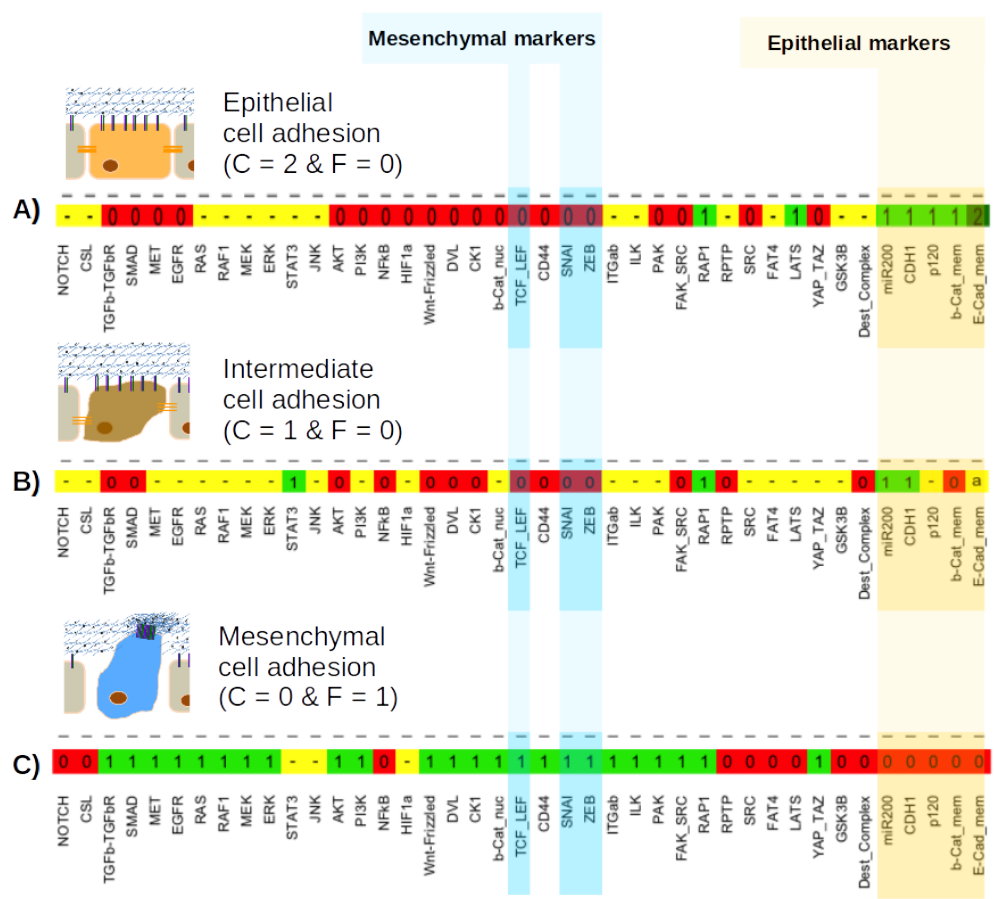
**Figure 20. Combinations of the model outputs (adhesion properties) and compatible phenotypes.** The combinations that are recapitulated in the model computed stable states are indicated in yellow. The phenotypes that were not recapitulated (absent) in the stable states are indicated in grey. The model stable states were obtained as is described in section 2.3.1. The values of the model outputs are indicated below the arrows. Phenotypes were distributed according to the characteristic cell adhesion properties in [10,12,13] (see also section 2.3 in chapter 1 for description).

### 3.4. Recapitulated biological markers and patterns

Here, we further test the model by comparing the computed stable states of the adhesion properties of the recovered phenotypes with the reported markers of Epithelial and Mesenchymal cells (section 2.5.1). We obtained several distinct stable states for each combination of adhesion properties, discarding the model inputs (see section 2.3.2 for method and Figures S5-S6 for results). These were represented in one stable state to show only the conserved and variable activity of each intracellular network component (Figure 21).

For Epithelial adhesion properties (Figure 21A), our results show that high activity in Epithelial markers with basal activity of Mesenchymal markers are conserved. For Mesenchymal adhesion properties (Figure 21C), we obtained the opposite, where the nodes of Mesenchymal markers have high activity, and the nodes of the Epithelial markers basal activity. Together, these results show that the model is consistent with the reported Epithelial and Mesenchymal markers. For the Intermediate adhesion stable states (Figure 21B), the representative stable state shows the absence of Mesenchymal markers and the presence of the Epithelial markers miR200 and E-cadherin. Indeed, these features were experimentally observed during partial EMT of H1975 and MCF10A cell lines, supporting our results [34,104]. In addition, the activity of the nodes p120 and b-Cat\_mem in Figure 21B indicate that the binding of catenins with E-cadherin in the membrane is not a feature of intermediate adhesion phenotype, which is also in agreement with the observed status of p120 and  $\beta$ -catenin in *Drosophila* border cells [100]. Regarding the activity patterns of signalling components, the model stable states show that the joint activation of TGF $\beta$  and Wnt signalling is exclusive of Mesenchymal phenotypes Figure 21. This is in agreement with the pattern already observed in Mesenchymal cell lines of normal hepatocytes and

hepatocarcinomas (see sections 5.3.5 in Supplemental text for full comparison with stable states) [31].



**Figure 21. Representation of the model stable states of cell adhesion properties of Epithelial (A), Mesenchymal (B), and intermediate (C) phenotypes.** The stable states were computed as described in section 2.3.2. Red indicate basal activity, green indicates high (Boolean nodes) or intermediate activity (Multi-valued nodes), dark green indicates high activity (Multi-valued nodes). Yellow indicates nodes with variable activity and the remaining colours conserved activity. The symbol (-) in nodes in yellow indicate that all values are identified in the stable states. The letters in yellow indicate exceptions for values 0 (a), 1 (b) and 2 (c). The Epithelial markers are indicated in light orange and the Mesenchymal markers in light blue (see section 2.5.1 for references on the biological markers).

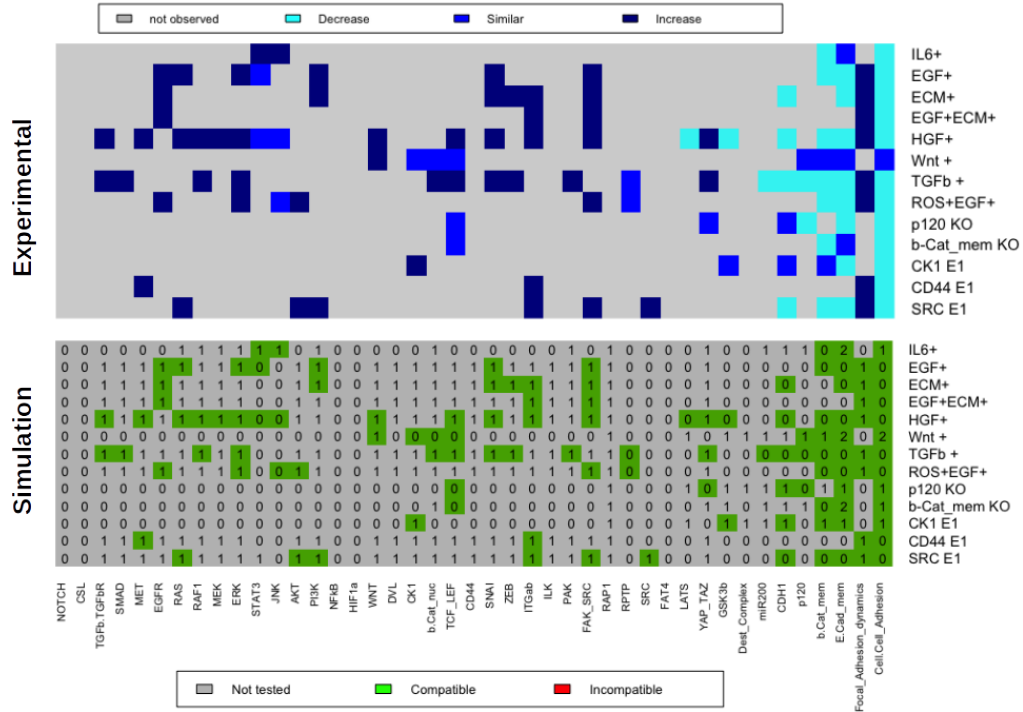
As expected, the activation of MAPK, AKT and integrin signalling are a characteristic of the Mesenchymal phenotype (Figure 21C) because they participate in the mechanisms of EMT initiation and progression [25]. Notch signalling and NFkB are not activated in the Mesenchymal stable states, which are in agreement with the idea that the enhanced activity is exclusive of cancer cells [25,91]. EMT associated activation of STAT3 and JNK is not always active in Mesenchymal stable states, indicating that these regulators are not crucial for EMT (see Figure S6). Interestingly, the model stable states (Figure 21A and Figure 21B) show that MAPK signalling, Notch signalling, STAT3, JNK and ILK activation by integrins are compatible with Epithelial and Intermediate phenotypes, indicating that this would depend on the microenvironment (see also Figures S5 and S7). Specifically for the stable states of Intermediate phenotypes (Figure 21B), our model shows that STAT3 is always activated, which is consistent with the observed STAT3 activation in *Drosophila* border cells during migration [107]. Moreover, our results indicate that RPTP, FAT4 and the Hippo signalling component LATS are basal in the Mesenchymal adhesion phenotype, which indicates incompatibility with this phenotype (Figure 21B). On the other hand, the stable states represented in Figure 21A and Figure 21C indicate that these molecular components are compatible with Epithelial and Intermediate phenotypes. This result further indicates that the activation degree of FAT4 and RPTP are characteristic features that distinguish Mesenchymal from Epithelial and Intermediate phenotypes.

### **3.5. Recapitulated experiments on Epithelial cell lines**

In this section, we present the recapitulated effects on cell adhesion properties and molecular activity observed in experiments on Epithelial cells by our model (for details see section 2.5.2). Except for neighbouring cell-cell contact signals

(NOTCH, RPTPL and FAT4L), we tested the individual effects of all remaining microenvironment signals in the model and/or in combination with another signal (Figure 22). With this data, we tested the model capacity to recapitulate the observed changes during EMT of Epithelial cells exposed to growth factors (EGF+, HGF+, and TGFb+) and the increase of ECM stiffness (ECM+). We also tested observed changes during EMT due to the overexpression of active SRC (SRC E1) and CD44 (CD44 E1). Moreover, we tested the model capacity to recapitulate experiments with a reported partial reduction in cell-cell adhesion in Epithelial cells due to IL6 exposure (IL6+), CK1 overexpression (CK1 E1), p120 knockout (120 KO), and the mutational effect that prevented the  $\beta$ -catenin/E-cadherin binding (b-Cat\_mem KO). In addition, we included an experiment, in which no changes in cell-cell adhesion was observed due to exogenous additions of Wnt to MDCK Epithelial cell lines (Wnt+). We simulated these 13 experiments (details in sections 2.4.1 and 2.4.2) and compared the resulting stable states with the collected observations of cell adhesion properties and molecular activity in these experiments (details in section 2.5.3). In these simulations, we assumed that the initial activity of Epithelial cells is described by the model stable state obtained with no inputs from the microenvironment and no mutations (see Table S4 for conditions used in simulations). we also assume no other signal from the microenvironment was present, except for the ones reported. This is reasonable to assume because we used controlled experiments to test the model. Each simulation resulted in a unique stable state, indicating a deterministic outcome (Figure 22). The results in Figure 22 show that in all experiments, the activity levels of the nodes resulted from simulation was compatible with the observations from experiments (no contradictory result was obtained). Thus, our results indicate that the model was able to recapitulate all 13 experiments. Previous model versions, which did not include all interactions of the model failed to

recapitulate some experiments (data not included in this thesis). For example, a version where the inhibition of free SRC by FAK/SRC complex was not accounted, resulted in activity levels of STAT3 incompatible with the experimentally observed in the experiments with EGF+ and HGF+ (data not shown).



**Figure 22. Comparison of the observations of experiments on Epithelial cell lines (top) with simulation results (bottom).** Experimental observations collected from published experiments on epithelial cell lines described in Table 5. Exogenous additions of microenvironment components are indicated by the symbol +, knockouts of intracellular network components by KO and overexpression/overactivation by the letter E. The value after the letter E indicate an ectopic activation at that value. For the experimental observations (**top**), the qualitative changes in activity are indicated by dark blue (increase), blue (no change) and cyan (decrease). For simulations (**bottom**), the value 0 means basal activity, value 1 intermediate (Multi-valued variables) or high activity (Boolean variables), and value 2 high activity (Multi-valued variables). Here, the nodes coloured with green indicates compatibility with experimental data and red incompatibility. Grey indicates predicted the activity of the nodes (not tested/observed experimentally). Simulations of experiments were carried in GINSim as described in sections 2.4.1 and 2.4.2. Parameters used in simulations are available in table S4.

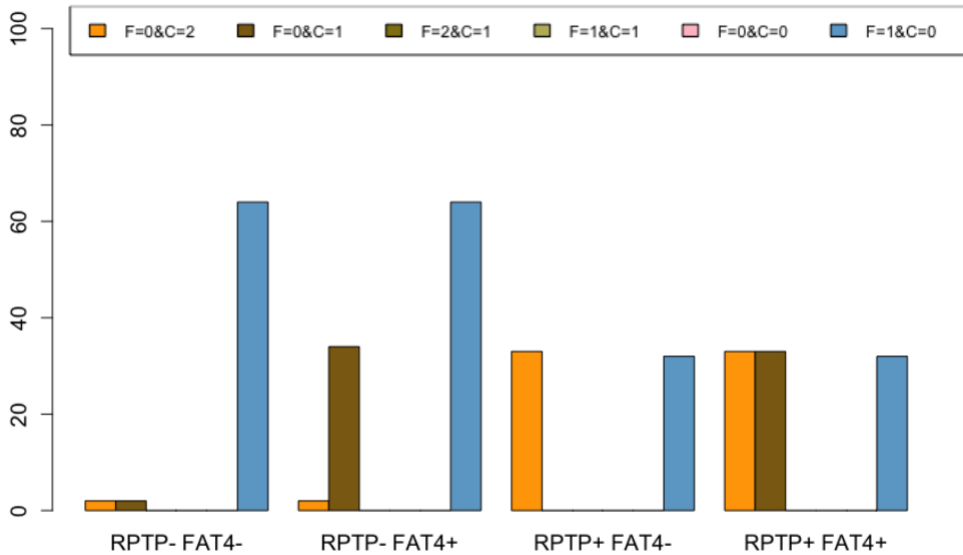


Concerning the quality of each experiment used to test the model, the activity observations covered between 12% to 51% of the nodes in the stable states. Together these experiments formed a data set that was able to cover 78% of the model intracellular components (Figure S8). Thus, we considered that this set of experimental data is a reasonable sample to test our model. With this data, we tested at least once, a component of all signalling pathways, except for the Notch signalling. Some nodes were tested more frequently than others, including the ones that belong to cell adhesion properties (model outputs), components of the E-cadherin/catenins complex, SNAIL, FAK-SRC complex, and ERK. Moreover, in some experiments we obtained more than one report for these nodes which further supported the confidence on the experimental observation (Figure S8). Together, this gives us higher degree of confidence on the model predictions for the above mentioned nodes, in particular, the ones related to cell adhesion properties (Cell-cell Adhesion and Focal Adhesion Dynamics).

### **3.6. Phenotypes compatibility with the microenvironment**

To Identify the microenvironment conditions compatible with phenotypes (model outputs) and to explore the role of cell-cell contacts, we analysed the model inputs combinations for each combination of RPTPL and FAT4L signals on the model stable states (section 2.3.4). The results of the analysis are available in Table S6 and the phenotypic effects of cell-cell contacts are highlighted in Figure 23. Our results show that the RPTPL signal substantially increases the number of microenvironments compatible with the Epithelial adhesion phenotype ( $F = 0$  &  $C = 2$ ) in comparison with the absence of these cell-cell contact signals. These results further indicate that the RPTPL signal ( $RPTPL = 1$ ) makes the Epithelial adhesion phenotype compatible with all possible combinations of signals that stimulate EMT (ECM, EGF, HGF and DELTA). In addition, these conditions are

only compatible with basal levels of ROS (value 0). This is expected because the RPTP (target of RPTPL) is directly inhibited by oxidation due to ROS, which in turn was accounted in our model by a direct regulatory interaction [108,109].



**Figure 23. The number of microenvironments compatible with adhesion phenotypes for different combinations of RPTP and FAT4 ligands.** The x-axis corresponds to the combinations of the model input values for RPTPL and FAT4L, where the presence is indicated by the symbol + (value 1) and the absence by the symbol - (value 0). The y-axis corresponds to the computed number of remaining model input combinations compatible with the model outputs (described in section 2.2.3). The remaining inputs are composed of the values of the nodes IL6, ROS, ECM, EGF, HGF, and DELTA (total of 64 combinations). The model output Focal Adhesion Dynamics is indicated by the letter F and the Cell-Cell Adhesion indicated by the letter C.

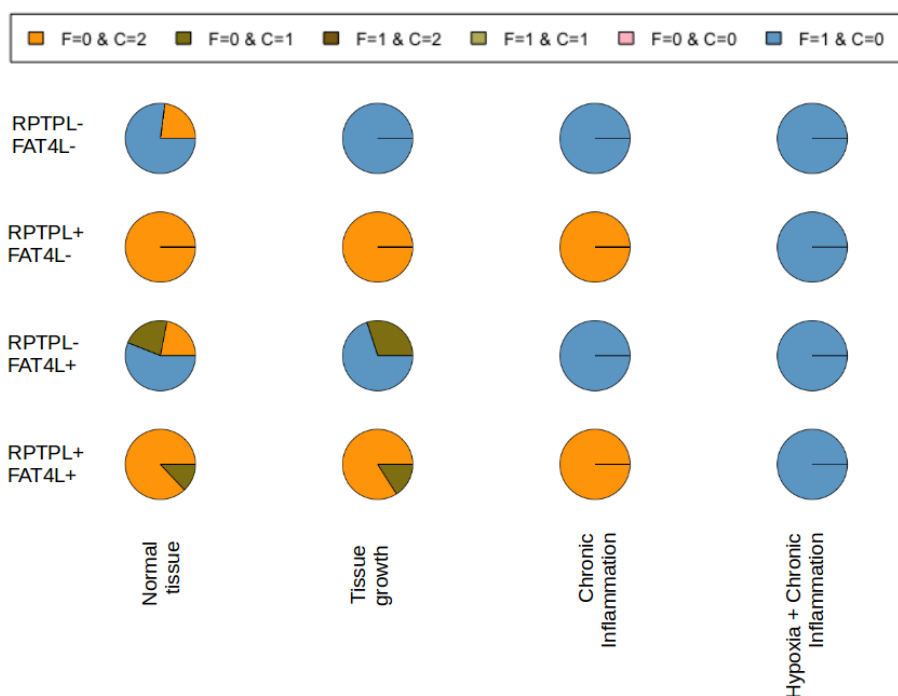
On the other hand, our results show that FAT4L signal (FAT4L = 1) makes the Intermediate adhesion phenotype (F = 0 & C = 1) compatible with all EMT inducing signals (ECM, EGF, HGF and DELTA). In addition, this effect is also compatible with the activation of RPTP by their ligands. We also obtained alternative conditions compatible with the Intermediate adhesion phenotype

which depended on the IL6 in the microenvironment ( $IL6 = 1$ ). For this condition, all EMT signals (ECM, EGF, HGF, and DELTA) were basal, together with the basal level of RPTPL signal. For the Mesenchymal adhesion ( $F = 1$  &  $C = 0$ ), our results indicate that this phenotype is compatible with all combinations of model inputs (microenvironments), except for the conditions where RPTPL signals are high and ROS levels are basal ( $RPTPL = 1$  and  $ROS = 0$ ). These results further indicate that the Mesenchymal adhesion is compatible with all remaining conditions in the absence of cell-cell contact signals that activate RPTP. Moreover, we identified conditions that resulted in both Epithelial and Mesenchymal adhesion phenotypes, indicating possible multistability. For the Intermediate and Mesenchymal adhesion phenotypes, the results also indicate a possible multistability for the conditions where FAT4L equals to 1. Thus, in these cases, the reachability of the adhesion phenotypes would depend on the initial conditions and the model dynamics.

### **3.7. The reachability of adhesion phenotypes**

To further explore the model dynamics and the effects of cell-cell contact signals on cell adhesion, we analysed the reachability of the model phenotypes in relevant physiological scenarios. Here, we further intended to reinforce the reachability of phenotypes by estimating their reachability probabilities, fixing the microenvironment signals (model inputs) and starting from all possible initial intracellular activity status (all state space). We chose to analyse the model in 3 microenvironment conditions that represent physiological conditions in Epithelial tissues and one prototypic tumour microenvironment (see Table S7 for conditions and references). For each of these conditions, we estimated the probabilities of reaching the adhesion phenotypes for combinations of RPTPL and FAT4L signals, starting from all model state space (section 2.4.2). The reachability

probabilities of the analysed physiological scenarios are presented in Figure 24. The obtained probabilities demonstrated that the RPTPL signal determines the reachability of the Epithelial adhesion stable states ( $F = 0$  &  $C = 2$ ) in normal, tissue growth and chronic inflammation microenvironment conditions.



**Figure 24. Reachability probabilities of cell adhesion phenotypes in physiologic scenarios.** Probabilities were estimated by model simulation in GINsim using the MonteCarlo algorithm (section 2.4.3). High activation (value 1) of RPTP and FAT4 by their ligands (RPTPL and FAT4L) are indicated with a + symbol and their basal activity (value 0) by a – symbol. For simulations, we fixed the remaining inputs (IL6, ROS, ECM, EGF, HGF) to the values according to the Table S7 and started from all state space.

On the other hand, these results also shows that the Mesenchymal adhesion phenotype ( $F = 1$  &  $C = 0$ ) is the only reachable phenotype when hypoxia ( $ROS = 1$ ) is combined with chronic inflammation signals ( $IL6 = ECM = HGF = 1$ ). This is because ROS directly inhibits RPTP making the signals in chronic inflammation conditions to determine the Mesenchymal phenotype outcome whatever the initial

state. This result is in agreement with an enhanced stimulation of EMT observed in tumour microenvironments composed by hypoxia with the chronic inflammation conditions [22,23].

In addition, the simulation of the hypoxia effect alone was not able to trigger EMT, and the combination with ECM or HGF resulted in the Mesenchymal phenotype (data not shown). This further indicates a synergistic effect of the combination of hypoxia with EMT driving signals such as growth factors or ECM for the stability of the Mesenchymal phenotype. For chronic inflammation conditions, our results show that cell adhesion stability is only affected by the RPTPL signal (no effects for FAT4L). Interestingly, when RPTPL and/or FAT4L signals are active, multiple reachable phenotypes including Epithelial, Mesenchymal, and Intermediate were obtained in normal and tissue growth microenvironments. This indicates that in these physiological conditions, cell adhesion depends on the initial state, i.e. there is a set of initial activity states that conditioned the reachability of the adhesion phenotypes. We also analysed the effect of DELTA signal in combination with RPTPL and FAT4L signals, obtaining similar results. In this analysis, we only identified a significant effect of DELTA in normal conditions and in basal RPTPL and basal FAT4L signals, where it resulted in the Mesenchymal phenotype with a probability of 1 (data not shown). This effect is in agreement with the reported effect of DELTA in driving EMT through the activation of Notch signalling [91,110].

### **3.8. The effect of model perturbations on phenotypes**

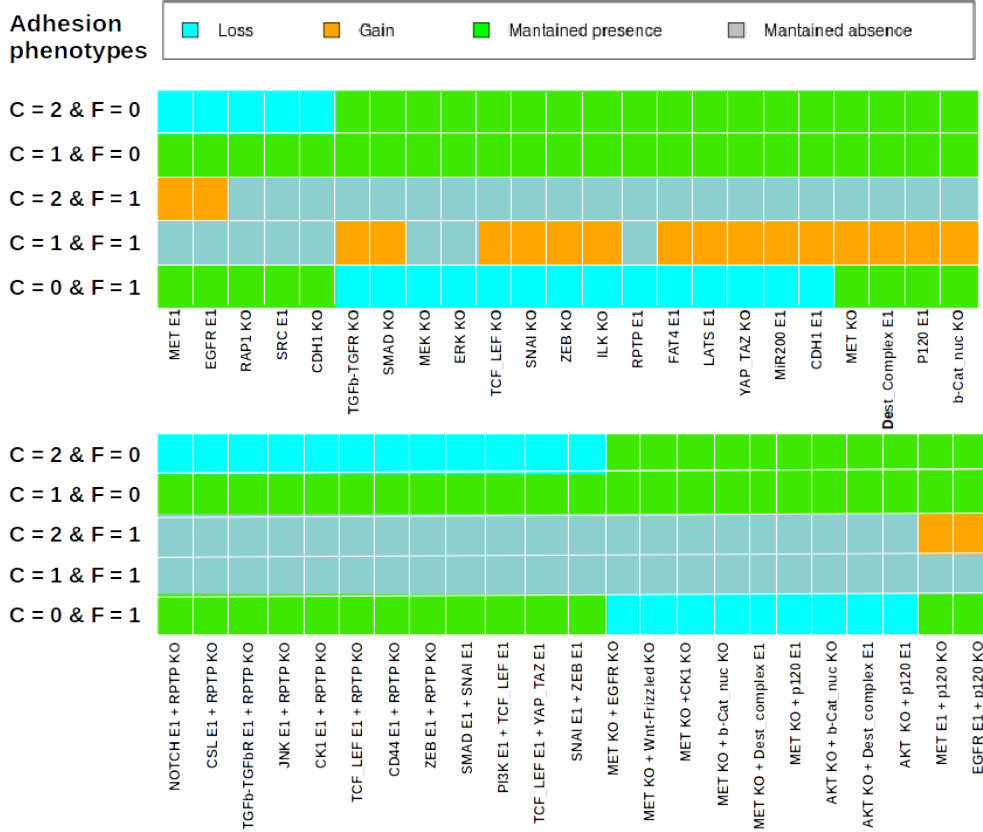
To identify critical alterations for the stability of Epithelial, Mesenchymal and Hybrid adhesion phenotypes, we generated perturbations on the model components and compared the presence of phenotypes in the stable states with the ones from the unperturbed model (see section 2.3.1 and 3.3.4 for methods).

Here, we systematically analysed the effects of all single and double perturbations on model components by fixing their values to be constant mimicking knockouts (KO) or ectopic activations (E1 and E2). In this analysis, we identified 23 single and 24 double perturbations that resulted in a gain and/or a loss of an adhesion phenotype (Figure 25). For single perturbations, we obtained 5 perturbations that prevented the Epithelial adhesion phenotype ( $C = 2$  &  $F = 0$ ) keeping the Mesenchymal ( $C = 0$  &  $F = 1$ ) and the Intermediate phenotypes ( $C = 1$  &  $F = 0$ ). Among them, only the perturbations of SRC and MET ectopic activation (SRC E1 and MET E1) correlate with higher expression of these proteins in carcinomas, indicating that they could be relevant in the context of cancer invasion (Table S8). SRC overexpression and its activation have been reported in a variety of human cancers including breast and colon cancer, which supports the relevance of this perturbation [111]. Interestingly, MET and EGFR ectopic activation were able to stabilize the Hybrid adhesion phenotype ( $C=2$  &  $F=1$ ), which indicates that these RTKs are critical for the stability of this phenotype.

On the other hand, we found 14 single perturbations that were capable of collapsing the Mesenchymal adhesion phenotype keeping both Epithelial and Intermediate adhesion phenotypes. These perturbations indicate that targeting components of TGF $\beta$ , Wnt, MAPK, Hippo signalling, and transcriptional regulators of E-cadherin (CDH1, miR200, ZEB and SNAI) leads to a change in phenotype stability. Most of these perturbations resulted in the stability of the Hybrid adhesion phenotype with an intermediate cell-cell adhesion ( $F=1$  &  $C=1$ ). Among them, the perturbations in E-cadherin regulators and the TGF $\beta$  signalling are in agreement with the predictions from published models [8,20,35]. Interestingly, the comparison with the expression in carcinomas further indicated that the perturbations in the Hippo signalling components, including FAT4 ectopic

activation, are overexpressed in several carcinomas indicating that they could be relevant for cancer invasion (Table S8). Although most of these KO perturbations do not correlate with the expression observed in carcinomas, some reflect inhibitory effects of drugs used in cancer therapy such as the SMAD inhibitors [35,112].

In contrast, the results indicate that the inhibition of MAPK signalling by its components (MEK and ERK) or the ectopic activation of RPTP were able to prevent the Mesenchymal phenotype without stabilizing a Hybrid phenotype. We also obtained 2 double perturbations that resulted in the stabilization of a cell adhesion combination (C=2 & F=1) compatible with the Hybrid phenotype. However, they do not correlate with the observed expression in carcinomas (Table S9). For the destabilization of Epithelial adhesion, we obtained several combinations of perturbations that resulted in a synergistic effect, where 3 correlate with the expression in carcinomas (Table S9). Two of them were the KO of RPTP in combination with NOTCH or TCF/LEF ectopic activation, which further indicates that RPTP activation is critical for cell adhesion. The other was the combination of PI3K with TCF/LEF ectopic activation, which have both high expression in carcinomas. Finally, we also identified several synergistic combinations that prevented the Mesenchymal adhesion phenotype. These perturbations indicated that the inhibition of AKT or MET in combination with the inhibition of particular Wnt signalling components (e.g. CK1) resulted in the suppression of the Mesenchymal phenotype without resulting in a Hybrid phenotype. These predictions inversely correlate with the expression in carcinomas (Table S9), which places them as possible combinatorial drug targets to prevent EMT.



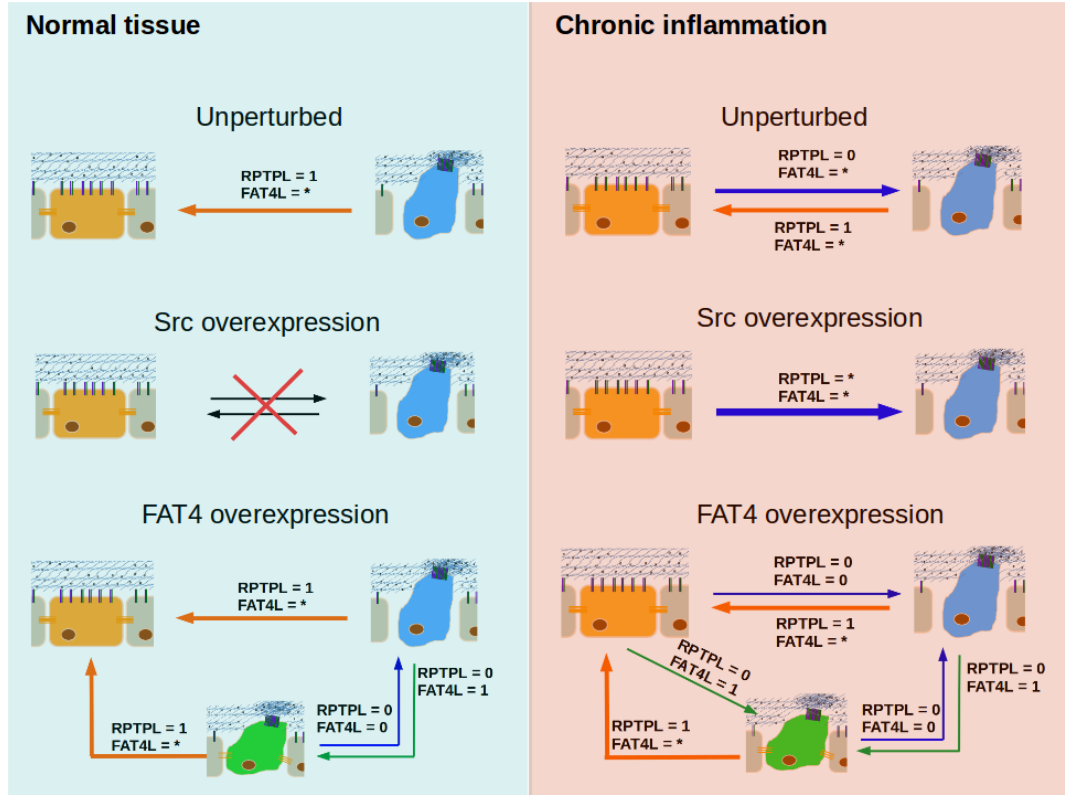
**Figure 25. Effects of single and double perturbations on the model adhesion phenotypes.** Single and double perturbation on model components are indicated by the nodes with a KO, E1, or E2, indicating that the values of the nodes were fixed to 0, 1 or 2. Model phenotypes are represented by the values of focal adhesion dynamics (F) and cell-cell adhesion (C). The effects of perturbations are indicated by the colours that indicate a gain or a loss of an adhesion phenotype in comparison to the unperturbed model. The perturbations that did not cause changes in the adhesion phenotypes are indicated by the colours green (present in both perturbation and unperturbed model) and grey (not present in both perturbation and unperturbed model).

### 3.9. The impact of SRC and FAT4 overexpression

To further understand the effects of cell-cell contact signals RPTPL and FAT4L in the cancer context, we compared the switching behaviour between phenotypes



due to these signals, accounting for two critical alterations observed in carcinomas. As critical alterations, we focused on the overexpression of SRC and FAT4 previously demonstrated to affect the stability of Epithelial and Hybrid phenotypes (section 3.8). To be more realistic, we mimic the overexpression by perturbations that made SRC and FAT4 insensitive to their inhibitors instead of an ectopic activation (section 3.4.1). The reasons are based on the dependence of FAT4 activation by its ligand and SRC activation by EGFR signalling [113,114]. Here, we compared the effects of RPTPL and FAT4L signals on the model phenotypes switching behaviour with and without these perturbations in two physiological scenarios (Figure 26). The unperturbed model showed that under normal Epithelial tissue and chronic inflammation conditions, the RPTPL signals activate the phenotype switching from Mesenchymal to Epithelial cell adhesion (MET), regardless of any other cell-cell contact signals. This switching was not possible with the SRC overexpression perturbation but was still present in the FAT4 overexpression. This is explained by the direct inhibition of RPTP due to free SRC activation, which is only active in the Mesenchymal stable states in the case of the SRC overexpression perturbation (Figure S11). This free SRC activation in the stable states is due to the removal of the inhibitory effect of FAK/SRC complex (interaction 124, Table S2). This further indicates that the proportions of FAK and SRC in the cell may be critical for the regulation of RPTP [115,116]. Under the SRC perturbation, the model predicts that the Epithelial phenotype was only stable in normal tissue conditions (no EMT signals), whereas it switches to Mesenchymal adhesion phenotype regardless of the FAT4L and/or RPTPL signals in chronic inflammation conditions (EMT signal). These results indicate that the overexpression of SRC leads to an irreversible switching behaviour towards the Mesenchymal phenotype. This further supports the idea of the potential role for SRC overexpression in cancer invasion [111].



**Figure 26. Comparison of the model phenotype switching behaviours due to cell-cell contact signals with and without perturbations that mimic SRC and FAT4 overexpression.** Adhesion phenotypes are represented by the correspondent Epithelial (orange), Mesenchymal (blue) and Hybrid (green) illustrations (see figure Figure 20 in section 3.3. for adhesion combinations). Arrows indicate switching between phenotypes and the input conditions for RPTPL and FAT4L are indicated on top of the arrows, where \* indicates all values are compatible. SRC overexpression was simulated by the perturbation that removed the inhibition effect of FAK\_SRC (SRC[FAK\_SRC@0] in Ginsim) and FAT4 overexpression by the perturbation that removed the inhibition effect by ERK (FAT4[ERK@0] in Ginsim). Simulations performed in GINsim with methods described in sections 3.4.2 and 3.4.3. Simulations started from the Epithelial, Mesenchymal or Hybrid stable states correspondent to the previously obtained stable state. The initial stable states were the ones obtained for normal tissue conditions. Simulations were run for all possible combinations of RPTPL and FAT4L values, fixing the remaining inputs as described for normal epithelial tissues or chronic inflammation (Table S6). Stable states from simulations are shown in Figures S10 to S12.

For FAT4 overexpression perturbation, our results indicate that this perturbation results in the stability of the Hybrid phenotype under the FAT4L signal ( $FAT4 = 1$ ) with basal RPTPL signals ( $RPTP = 0$ ). The resulting stable states of the Hybrid phenotypes were mostly identical to the Intermediate phenotype with intermediate E-cadherin levels, except for the activation of PAK, FAK-SRC, ERK and PI3K (Figure S7 and Figure S12). These additional activations explain the high focal adhesion dynamics in stable states because they are directly involved in the dynamics of focal adhesion (interactions 34, 42, and 44, Table S2). These activations were not compatible with FAT4 activation in the unperturbed model because FAT4 is inhibited by ERK (interaction 41, Table S2). The model also indicates that this perturbation leads to a reversible switching behaviour between Mesenchymal and Hybrid phenotypes, which only depends on the FAT4L signal strength. This is in agreement with the observation of both types of cells in migrating carcinomas [11]. For the switching from Hybrid to Epithelial phenotypes, the model showed that this behaviour exclusively depends on the RPTPL signal strength ( $RPTPL = 1$ ) in both physiological scenarios. Interestingly, the model dynamics showed that in normal Epithelial tissue conditions the Hybrid phenotype can only be formed from an initial Mesenchymal state containing and the FAT4 perturbation.

On the other hand, in chronic inflammation conditions, the switching from Epithelial to Hybrid is already possible in high FAT4L and basal RPTPL signals. This indicates that an EMT signal is required for driving the switching to the Hybrid phenotype in the case of FAT4 overexpression. We also found that IL6 signal in combination with any the signals that drive EMT was necessary to drive the switching from Epithelial to Hybrid phenotype. We also analysed the combinations of RPTPL and FAT4L with the DELTA signals, where no striking differences were found between the unperturbed and perturbations in terms of

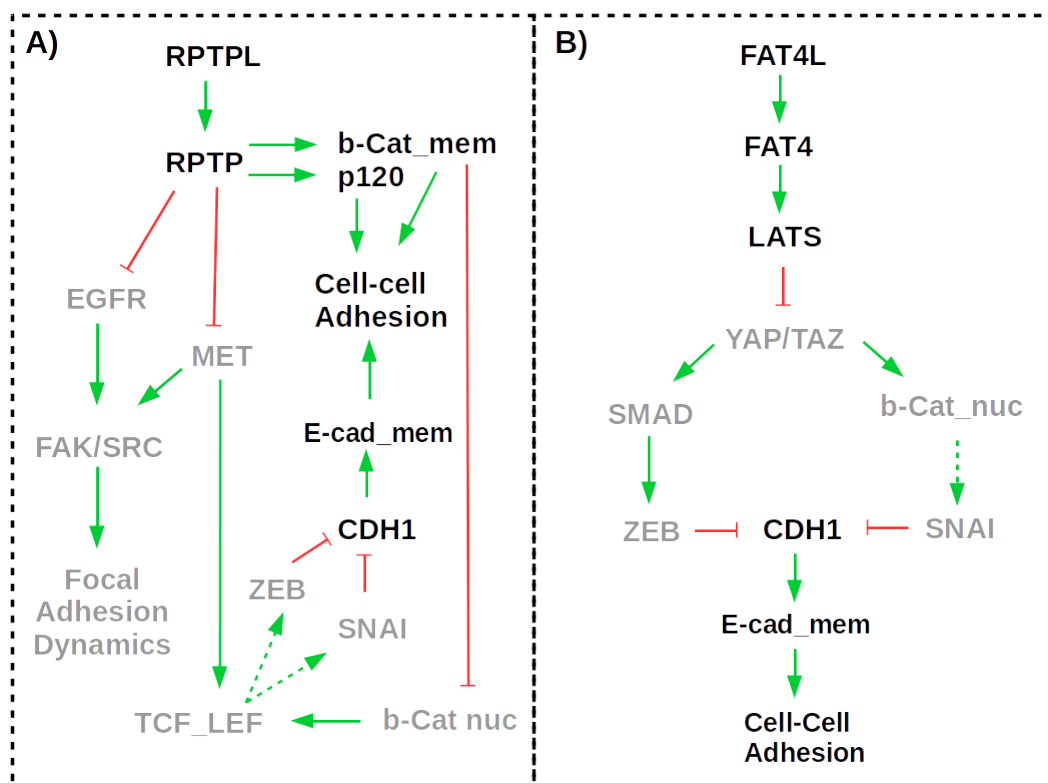
the switching between Epithelial, Mesenchymal and Hybrid phenotypes (data not shown). However, in some conditions, the DELTA signals were able to stabilize the Intermediate phenotype in the normal and FAT4 overexpression perturbation. Some of these conditions were compatible with the Hybrid phenotype (RPTPL=0 and FAT4L =1), indicating that DELTA signals may be involved in the formation of a cluster of cells composed of Hybrid and Intermediate phenotypes.

### **3.10. The mechanism of RPTP and FAT4 on cell adhesion**

To explain how RPTPL and FAT4L signals control cell adhesion properties under an EMT signal, we analysed the impact of several perturbations that knockout regulatory effects (model interactions) downstream from these signals. For this purpose, we analysed the reachability of the adhesion phenotypes under chronic inflammation conditions for individual and combination of perturbations (similar to the analysis in section 3.7). The results of the reachability of phenotypes (Tables S10 and S11) allowed us to identify the effect of regulatory interactions, discard interactions, and reconstruct the regulatory mechanism by which RPTPL and FAT4L signals control cell adhesion properties (Figure 27). For the effect of the RPTPL signal, our results demonstrated that RPTP regulatory effects on EGFR, MET,  $\beta$ -catenin (b-Cat\_mem), and p120 are required to ensure an absolute control of cell adhesion properties whatever the initial state (Figure 27A). The reachability analysis showed that the inhibition of MET together with activation of p120 and  $\beta$ -catenin is critical to ensure Epithelial cell-cell adhesion in chronic inflammation conditions. This was evident in the triple perturbation knockout of these interactions, which resulted only in the Mesenchymal cell adhesion phenotype (Table S10). Here, the RPTP activation of the membrane forms of p120 and  $\beta$ -catenin are both required to complement the effect of E-cadherin on

cell-cell adhesion strength leading to the conditions for high Epithelial cell-cell adhesion [117,118].

For MET inhibition, RPTP plays a key role to effectively prevent the inhibition of E-cadherin expression (CDH1) by SNAI and ZEB via TCF/LEF activation of Wnt and TGF $\beta$  signalling, otherwise would completely abolish Epithelial adhesion phenotype (Figure 27A, Table S10).



**Figure 27. Critical interactions that are involved in the control the cell adhesion properties under chronic inflammation conditions with RPTPL (A) and FAT4L (B) signals.** Regulatory components represented by the nodes of the model (see Table S1 for full names). Components in grey indicate basal activity and in black high or intermediate activity. Arrows in green indicate activations and red inhibitions. Continuous arrows indicate a direct interaction between two nodes in the model. Dashed arrows indicate that the interaction effect may be through multiple paths.

On the other hand, the joint inhibition of MET and EGFR is required to ensure an absolute control over the inhibition of the Mesenchymal focal adhesion dynamics via the inhibition of FAK/SRC complex. This was evident in the stabilization of high focal adhesion dynamics in the ectopic activation of FAK/SRC complex by MET or EGFR (Table S10). In addition, our results also indicate that ineffective control of RPTP over EGFR or MET leads to the stabilization of Hybrid adhesion phenotypes, and the ineffective control of RPTP over catenins (p120 and  $\beta$ -catenin) resulted in loss of focal adhesion dynamics that could result in an amoeboid-like phenotype (Table S10).

For the control of cell adhesion properties by FAT4L signals, our analysis showed that the activation of Hippo signalling component LATS and subsequent inhibition of the nuclear translocation of YAP and TAZ (YAP\_TAZ ) is critical for the stability of the Hybrid adhesion phenotype (Figure 27B, Table S11). This is demonstrated by the reachability of the adhesion phenotypes with perturbations that removed the effect of these interactions, which resulted in the Mesenchymal phenotype instead of the Hybrid (Table S11). The high focal adhesion dynamics on these phenotypes is always maintained under chronic inflammation conditions because this property is controlled by regulators that do not interact with the Hippo signalling components. On the other hand, the Hippo signalling perturbations affected only cell-cell adhesion, in particular through the YAP/TAZ participation in the activation of nuclear  $\beta$ -catenin (b-Cat\_nuc) and TGF $\beta$  signalling component SMAD. These two interactions are both required for the stabilization of the cell-cell adhesion correspondent of the Hybrid adhesion phenotype, where the effect of the YAP/TAZ interaction with DVL was neglectable. This was demonstrated by the reachability of both Mesenchymal and Hybrid adhesion phenotypes with perturbations on these two interactions (Table S11). Moreover, the inhibition of SMAD and nuclear  $\beta$ -catenin further explains the stability of the Hybrid

intermediate cell-cell adhesion by having a downstream effect, repressing E-cadherin expression through ZEB and SNAI activation (Figure 27B).

#### **4. Discussion**

In this study, we proposed a logical model that focused on the regulation of cell adhesion properties during EMT. This model was developed to provide an insight on the conditions that control the switching behaviour between Epithelial, Mesenchymal and Hybrid phenotypes. Here, we have accounted for the regulation of E-cadherin mediated cell-cell adhesion and the focal adhesion dynamics to provide a simple way to define Epithelial, Mesenchymal and Hybrid phenotypes. This allowed us to better understand the role of the microenvironment, in particular, the role of neighbouring cells signals towards invasion and metastasis of carcinomas. Our systems approach allowed us to account for the complex regulation of EMT, recapitulate several reported observations (model validation) and predict the conditions that drive the switching between Epithelial, Mesenchymal and Hybrid phenotypes. In this work, we have further focused on two cell-cell contact signals that activate the RPTP and FAT4 receptors as neighbouring cells signals. Moreover, we explored the impact of some reported mutations in carcinomas on the switching behaviours.

Our model was able to recapitulate the Epithelial and Mesenchymal phenotypes on the stable states, together with their reported markers and most of the expected changes in the activity of signalling pathways during EMT (Figure 21). From the analysis of the Epithelial and Mesenchymal stable states patterns, we have learned that characteristic signalling activation of EMT depend on the microenvironment, where some reported signalling activation in EMT may be activated without resulting in a phenotype change (Figure 21). In addition, the model was also able to recapitulate the cell adhesion properties and reported

molecular features of the border cells migration during developmental process of oogenesis by the Intermediate adhesion phenotype (Figure 21) [14,34,100]. Accounting for the microenvironment signals analysed here, our model predicted that the Hybrid and Amoeboid-like phenotypes were not stable under any combination of microenvironment signals (Figure 20). Although this was part of the model validation, it suggests that the stability of these phenotypes depends on mutation acquisition during cancer progression that causes regulatory alterations such as KOs or overexpression of critical genes [28,119]. To reinforce the confidence in our model predictions, we have tested the model capacity to recapitulate the effects of microenvironment signals and mutations on Epithelial cell lines (Table 5, Figure 22). Here, the model was able to recapitulate the observed adhesion changes and molecular activity during EMT triggered by growth factors, ECM stiffness and mutational effects such as SRC overactivation (Figure 22). The model also recapitulated partial adhesion changes caused by mutations or IL6 in epithelial cells (Figure 22). Most of the experiments used came from different cell lines which some were cancerous (Table 5), bringing noise and uncertainty around the specificity of the results. Logical models have been validated using either expression data or specific measurements of activity in particular cell lines [31,32,38,39]. Since most of our model components are signalling components, the expression data is not suitable to test our model because it does not directly reflect the activity degree [120,121]. On the other hand, controlled experiments that tested model components in the same cell line and within the same experimental method are not abundant in the literature. Nevertheless, we have covered a substantial amount of model components by aggregating observations from several cell lines and studies (Table 5), where the adhesion properties and its direct regulators were constantly challenged (Figure 22 and Figure S8). Thus, we considered the model to be able to generate



trustable predictions for generic Epithelial, Mesenchymal and Hybrid cell types, in particular, for their relationship with the microenvironment.

The model analysis revealed that cell-cell contact signals that activate RPTP were determinant for the control cell adhesion properties during EMT (sections 3.6, 3.7, and 3.9). Our model showed that RPTP activation was able to make the stable states that belong to the Epithelial adhesion phenotype compatible with the presence of signals in the microenvironment that are reported to trigger EMT (e.g growth factors, ECM stiffness and Delta-notch activation) (Figure 23 and Table S6). Interestingly, the reachability analysis further showed that RPTP activation by cell-cell contact signals resulted into the Epithelial phenotype in normal, tissue growth and chronic inflammation conditions (Figure 24). Together, these results suggest that cell-cell contact signals from neighbouring cells are capable of strongly activate RPTP could prevent EMT and trigger Mesenchymal-to-Epithelial Transition (MET). This constitutes a model prediction that explains the observed MET in cancer cell lines due to signals from co-cultured normal Epithelial cell lines (cell-cell contact stimulation) [65]. Our prediction also explains EMT in tissues that suffered massive cell death (loss of cell-cell contacts) under tissue repair signals, followed by MET during the process re-epithelialization (regain of cell-cell contacts by repopulation of cells) [102]. These RPTP effects are not described in the literature and suggest that they might be relevant in the cancer context. This role is supported by the observation of cell-cell dependent RPTPs such as RPTP- $\kappa$  and DEP-1 are often mutated and down-regulated in carcinomas [40,122]. Our analysis further showed that the control of cell adhesion by RPTP depend on the combined activation of the membrane forms of p120 and  $\beta$ -catenin, together with the inhibition of the EGFR and MET (Figure 27, Table S10).  $\beta$ -catenin and EGFR are specifically targeted by RPTP- $\kappa$ , whereas p120 and MET are strongly dephosphorylated by DEP-1 [40,123–125]. This

further suggests that both RPTPs are critical for the control of EMT, possibly through a combined effect with more RPTPs to ensure a stronger effect.

Importantly, our results also showed that high ROS levels in the microenvironment (e.g. generated under hypoxia) favours EMT by preventing the RPTP control mechanism (Table S6 and Figure 24). However, this required additional signals such as growth factors or ECM stiffness to trigger the switching towards Mesenchymal phenotype (Table S6). This suggests that the oxidative inhibition of RPTPs by ROS as it was observed for RPTP-k in keratinocytes may promote carcinomas invasion through the accumulation of Mesenchymal cells [108]. This effect was strikingly evident in the simulations of hypoxia with chronic inflammation conditions, suggesting that these signals in the tumour microenvironment may be enough to promote invasion of carcinomas, regardless of mutational acquisition (Figure 24). This explains the direct link between metastasis and the observed synergistic effect of combining hypoxia with high ECM stiffness in the tumour microenvironment [23]. Similar effects were also identified for the model perturbation that mimics the overexpression of SRC in carcinomas (sections 3.9 and 3.10). In this case, the perturbation allowed the activation of the free SRC form as well as the complexed with FAK to mimic the higher concentrations of SRC in cells (Figure S10 and S11). Consequently, the free SRC activation in the Mesenchymal stable states results in the inhibition of RPTP transcription in our model (interaction 108) [126]. This caused an irreversible switching behaviour that would favour the accumulation of Mesenchymal cells in comparison to the absence of a SRC overexpression under chronic inflammation conditions (Figure 26). This suggests that mutations that result in an increase of SRC concentration in carcinomas would favour invasion through gaining the capacity to overcome EMT control mechanism by RPTP [111,115]. This supports the potential role of SRC overexpression and activation

in cancer invasion, together with the observed down-regulation of RPTP-k in similar experimental conditions [111,126,127]. This finding is further supported by the observation of normal RPTP-k expression during EMT in keratinocytes and breast cell lines, under experimental conditions where SRC was not overexpressed [128,129]. Perturbation analysis of model components (section 3.8) further highlighted the relevance of RPTP in cancer by showing several perturbations that would favour the Mesenchymal phenotype only if RPTP is knockout (Figure 25). For example, Notch or TCF/LEF overactivation were two cases that we found to correlate with the expression levels in carcinomas (Table S9). Together, these results reinforce the idea that RPTP activation by neighbouring cells may be critical to prevent cancer invasion and metastasis through an uncontrollable EMT.

On the other hand, our model analysis showed that cell-cell contact signals that activate FAT4 receptor could stabilize Intermediate and Hybrid cell adhesion phenotypes in the presence of microenvironment signals that trigger EMT (Figure 23, Table S6, and Figure 26). This suggests that neighbouring cells signals such as Dcsh1 ligands involved in ureteric bud branching and skeletal morphogenesis could play a role for the stability of partial EMT phenotypes [43,113]. However, this is predicted by our model to be incompatible with RPTP activation by neighbouring cells signals, suggesting that deregulations of RPTP in cancer would favour the stability of partial EMT phenotypes (Table S6 and Figure 26). Thus, mutations in RPTP ligands, hypoxia or SRC overexpression are possible deregulations that could also play a role in stabilizing partial EMT phenotypes. Interestingly, our model predicted that the stabilization of a Hybrid adhesion phenotype was possible through a perturbation that reflects the overexpression of FAT4, i.e., a mutational alteration that removed the reported transcriptional control of FAT4 by ERK activation (Figure 26) [130]. This perturbation resulted in

the Hybrid adhesion with an intermediate cell-cell adhesion as the Intermediate phenotype, which is in agreement with the expected adhesion properties of the small clusters observed during cancer migration [14,15]. Despite the known role in border cells migration, the Intermediate adhesion phenotype could also play a role in the collective migration in cooperation with the Hybrid phenotype as follower cells (Hybrid + Intermediate cells) [14,100]. Therefore, our results suggest that overexpression of FAT4 could promote invasion of carcinomas through favouring collective cell migration. Indeed, the FAT4 expression is increased in most carcinomas, supporting that this alteration could be relevant for cancer invasion (Table S6). Strikingly, this prediction is in contradiction with the reported tumour suppressor role of FAT4 via activation of Hippo signalling, making this a new hypothetical role of FAT4 in cancer [41].

The stability of the Hybrid phenotype by FAT4 activation was explained with our model by a joint inhibition of SMAD and TCF/LEF due to YAP/TAZ inhibition by Hippo signalling activation (Figure 27, Table S11). In turn, this combined inhibition was required for successfully prevent the down-regulation of E-cadherin expression by their main transcriptional repressors (ZEB and SNAI family) under any EMT signal [27,131,132]. However, it requires additional EMT signals such as the ones from chronic inflammation to maintain integrin signalling towards activating high focal adhesion dynamics. The explored dynamics of phenotype switching under FAT4 overexpression perturbation further showed that signals from neighbouring cells conditioned the inter-conversion between Hybrid, Mesenchymal and Epithelial (Figure 26). Here, we have illustrated a tendency for invasion under chronic inflammation conditions with inactivation of RPTP (deregulation). Here, we also captured a tendency for colonization by the irreversible switching to the Epithelial phenotype in normal tissue conditions with RPTP reactivation by neighbouring cells (Figure 26). This recapitulates the

metastasis process and places neighbouring cells signals as playing a central role together with changes in the microenvironment between the primary and secondary sites of tumours [4,9,18]. Under these conditions, Intermediate phenotypes are also predicted to be possible further supporting the possibility of the formation of an invasive cluster of tumour cells (Figure 26, Table S6). In addition, we also identified that Delta-notch signalling could play a role in the interchange between Hybrid and Intermediate phenotypes in these invasive clusters (section 3.9). Comparing with the dynamics of inter-conversion of SRC overexpression, our results support the idea that metastasis would be further stimulated in a heterogeneous tumour containing cells with these individual mutations, stimulating both single (Mesenchymal) and collective tumour invasion (Hybrid + Intermediate) (Figure 26) [111].

Although we have not explored in depth all mutational effects, some mutational alterations were pinpointed to be relevant for the invasion of carcinomas (section 3.9). One was the cMET overactivation that allowed the stabilization of the Hybrid phenotype with high cell-cell adhesion (Figure 25). This cell adhesion is characteristic of a multicellular stand invasion, suggesting that cMET overactivation might play a role in this form of invasion [14,15]. Thus, keeping cMET activation controlled by RPTPs or drugs (inhibitors) is predicted to prevent this phenotype. In this analysis, we also identified that the combination of PI3K with TCF/LEF activation was critical for overcoming RPTP control and collapsing the Epithelial phenotype (Figure 25). Indeed, PI3K is found to be frequently overactive in carcinomas and correlated with aggressiveness, supporting its relevance for cancer invasion [28]. However, TCF/LEF is a transcription factor that depends on the activation of Wnt signalling pathways [133,134]. Interestingly, we also found that some KO perturbations in TGF $\beta$  and Wnt signalling components stabilized the Hybrid adhesion phenotype compatible with the

collective migration of small groups of cells. Among them, SMAD has been used as a drug target in cancer therapy, proven to be ineffective and predicted to stimulate partial EMT [35,112]. On the other hand, our model predicted that targeting ERK/MEK or the combination of AKT/MET with CK1 could be sufficient to prevent Mesenchymal phenotypes. These results support the usage of effective MAPK inhibitors as drugs to fight cancer [87,88]. Or in alternative, the idea of targeting AKT and MET in combination with other drugs to increase the efficiency [137,138]. Together, these results suggest that depending on the microenvironment, the usage of drugs in cancer therapy could result in alternative modes of invasion (Hybrid or Mesenchymal phenotypes).

Taken all together, our results showed that neighbouring cells signals in the tumour microenvironment play a role in metastasis by driving the interconversion between Epithelial, Mesenchymal and Hybrid phenotypes. RPTP activation by neighbouring cells is predicted to function as a natural mechanism to control EMT. Here, we showed that overcoming this control mechanism is possible through deregulations in the tumour microenvironment (e.g. hypoxia combined with chronic inflammation) or mutational deregulations (SRC overexpression), explaining invasion of carcinomas through EMT. Moreover, overexpression of FAT4 receptors is pinpointed by our model as a critical mutational deregulation that could promote collective cell migration of carcinomas by allowing the stability of Hybrid phenotypes.

## **5. Supplemental Information**

### **5.1. Supplemental text**

For most nodes, we only consider the post-translational regulation for the activation or inhibition of a component (see Tables S1 and S2). The reason is

because these nodes reflect proteins of signalling cascades, and we did not find evidence for transcriptional regulation. To make our analysis computationally feasible, we reduced the size of the network by collapsing mediators of regulatory processes such as adapter proteins, genes and RNA (assumed to be part of an interaction). In these cases, we added a direct arc between the regulator and target which bypassed mediators. Protein complexes that act as a scaffold or families of proteins/genes with redundant effects were condensed into a single node. We present some of these simplifications in the KEGG signalling pathways in Figures S1-S4.

#### **5.1.1. Interactions for the regulation of cell-cell adhesion**

The status of the E-cadherin/ $\beta$ -catenin complex on the adherens junctions was adopted as a proxy of cell-cell adhesion strength, neglecting other adhesion proteins [117,118,139]. The complex is represented by the nodes E-cad\_mem and b-Cat\_mem. We included two effects on cell-cell adhesion from these nodes, the adhesion caused by E-cadherin homophilic binding (interaction 16) and the strengthening of adhesion by  $\beta$ -catenin binding to E-cadherin (interaction 17) [117,140]. We neglected the effect of  $\alpha$ -catenin on cell-cell adhesion because it depends on  $\beta$ -catenin binding to E-cadherin [117,118]. For the regulation of E-cadherin/ $\beta$ -catenin complex, we included the effect of  $\delta$ -catenin (p120) on the stability of E-cadherin in the membrane (interaction 26) and the effect of RAP1 on the transport of E-cadherin to the membrane (interaction 25) [141,142]. To account for the mechanism of E-cadherin/ $\beta$ -catenin complex disassembly, we included the Inhibitory effects on ( $\delta/\beta$ )-catenin binding to E-cadherin by tyrosine kinases SRC, MET and CK1 (interactions 6-9 and 80-83, see also Figure S1 in section 6) [143]. We also included the effect of RPTP kappa and DEP1 on the activity of Receptor Tyrosine Kinases (interactions 33 and 71) and catenins

(interactions 4 and 80) [40,42]. Based on the redundancy of RPTP activity, we represented RPTP kappa and DEP1 by a single node (RPTP). Based on reports that showed RPTP kappa and DEP1 activation is enhanced in high cell confluency, we consider a cell contact activation of RPTP by a generic ligand (RPTPL) [44,45,144]. In this network model, the effects of intracellular phosphatases (e.g. PTPs and PTEN) were neglected assuming that the strength of tyrosine kinase activation is enough to become dominant over intracellular phosphatases [145,146]. This is reasonable to assume in the case of a strong signal to activate tyrosine kinases. We included the transcriptional repression of E-cadherin gene (CDH1) by ZEB and SNAI family of transcription factors (interactions 14 and 15, see also Figure S1 in section 6) [132,147,148]. This was coupled with the core EMT regulatory circuit composed by mutual inhibitions between ZEB/SNAI and miR200 family (interactions 72, 73 and 134) [34,104]. Canonic MAPK and TGF- $\beta$  signalling were included here because they mediate the activation of SNAI and ZEB, respectively (interactions 110, 116 and 133, see also Figure S1 in section 6) [25,34,80]. We accounted for the main growth factors in the microenvironment of Epithelial tissues able to drive EMT, the EGF and HGF [25,149,150]. Canonic Wnt signalling was considered because is involved in the autocrine regulation of TGF- $\beta$  signalling (interaction 130) and participates in the dissolution of E-cadherin/ $\beta$ -catenin complex through CK1 (interaction 81) [31,69,151]. In this pathway, we also included a self-activation on TCF\_LEF (interaction 129) to account for an enhanced activation of Wnt signalling due to AKT and MET effects reported in carcinoma cell lines [152,153].

### **5.1.2. Interactions for the regulation of focal adhesion dynamics**

For the focal adhesion dynamics we considered as direct regulators the integrin alpha5beta1 (ITGab), PAK1, and a complex composed by FAK and SRC (FAK-



SRC). Integrin  $\alpha 5 \beta 1$  was considered because ensures enhanced contractile forces necessary for tumour invasion (interaction 44) [97,154]. FAK-SRC and PAK were chosen because they both drive a dynamical formation and turnover of long F-actin stress fibres, necessary for high invasive cell motility (interaction 43 and 45) [13,155–158]. FAK isolated effect was not accounted because it is not correlated with invasive phenotype [140,159]. Because we chose FAK-SRC and PAK as direct regulators of focal adhesion dynamics, we neglected all downstream regulators such as the small GTPases Rac1 and RhoA (Figure S2 in section 6). For upstream regulation of focal adhesion dynamics, we accounted for several integrin and growth factor induced mechanisms (Figure S2 in section 6). We included the direct integrin activation by ECM stiffness (interaction 57) and growth factor induced activation by TGF- $\beta$ R and CD44 (interactions 54 and 55). The effects of ZEB and RAP1 on integrin overexpression and transport to the membrane were also accounted (interactions 53 and 56) [141,160]. PAK activation was found to be controlled by PI3K and FAK-SRC, thus we included these effects (interactions 85 and 86). For the mechanism of FAK-SRC complex formation and activation, we accounted for the regulatory effects of integrins, EGFR, cMET, free cytoplasmic form SRC and ERK (interactions 35-39) [94,140,155,161]. We assumed that free SRC in the cytoplasm is transient during focal adhesion formation. This was based on the observed re-localization of SRC from the cytoplasm to a FAK-SRC complex on focal adhesions in MDCK cells (normal cells) [115]. It is reasonable to consider that FAK-SRC binding would be enough to prevent the activation of SRC targets in the cytoplasm/nuclear locations (e.g. STAT3, interaction 124). To account for this, we included a self-activation on FAK-SRC node (interaction 40) and an inhibitory interaction from FAK-SRC to SRC (interaction 122). These interactions were necessary to account for a transient STAT3 activation due to SRC observed

in normal cells and constitutive activation when SRC is over-expressed [127]. We included several reported downstream effects from integrins and growth factor induced signalling. One was through the addition of ILK to account for integrin mediated activation of SMAD and AKT (interactions 3 and 109) [162,163]. Several FAK-SRC effects were included: The inhibition of membrane  $\beta$ -catenin (interaction 9); RAS mediated MAPK and JNK signalling activation (interactions 59 and 102); and PI3K mediated AKT signalling activation (interaction 89). Finally, we also included a putative inhibitory interaction from SRC to RPTP based on an observed negative correlation between RPTP kappa transcription and SRC activation in a SRC transformed MCF10A cell line (interaction 106) [126]. Since the mechanism of RPTP kappa inhibition is unknown and is only observed in conditions where SRC is overexpressed, we assumed that this effect is due to free SRC [126,128,129].

### **5.1.3. Interactions for the regulation of EMT and cancer invasion**

Notch, hypoxia and inflammatory IL6 signalling were included in the network based on reports that correlated these signals with EMT stimulation and tumour invasion [22,23,91]. In cancer cell lines, Notch signalling activates SNAI2 (interactions 20 and 111), stimulates EGFR overexpression (interactions 20 and 28), crosstalk with TGF- $\beta$  signalling and enhances invasion through NFkB activation (interactions 20 and 74). Inhibition of Notch by DVL (Wnt signalling) was found in MCDK cells, thus we included this effect (interaction 79). In the network, Notch activation was only considered through cell-cell contact signals by Delta-like ligands, represented by the node DELTA (interaction 78). We also considered ROS as a microenvironment signal because it can be generated by both inflammation and hypoxia in the tumour microenvironment [22,164]. This signal is important for triggering the expression of HIF1a (interaction 50), which

is involved in the stability of SNAI1 (interaction 115) and potentiates the effect of Notch in breast cancer migration [165,166]. STAT3 and NFkB also participate in HIF1a expression by preventing its turnover (included in interactions 49 and 51). In addition, we also included the effect of ROS in NFkB activation and RPTP kappa inhibition (interactions 76 and 105). For STAT3 activation, we include two interactions (123 and 124), to account for the IL6 and SRC induced constitutive activation of STAT3, which participates in stabilizing SNAI1 expression [127]. In IL6 signalling, we suppressed JAK (IL6 receptor) and add a direct link to JNK to simplify the STAT3 signalling based on evidence that showed JNK activation was required for a fully activated STAT3 transcriptional response [167,168]. On the other hand, JNK also promotes the disassembly of E-cadherin/ $\beta$ -catenin complex (interaction 6) and participates in SNAI activation (interaction 114) [169–171]. IL6 was considered as microenvironment signal neglecting its autocrine production because it is accumulated in tumours, and is thought to play a role in breast cancer invasion [22,172]. Hippo signalling was included in the network because hypothetically can counteract EMT based on its capacity to inhibit Wnt and TGF- $\beta$  signalling through LATS activation [95,173]. This causes cytoplasm retention of YAP and TAZ (interaction 132) which no longer participates in nuclear translocation of  $\beta$ -catenin and SMAD (interactions 10 and 107). Experimental evidence showed that LATS can be activated by either E-cadherin/ $\beta$ -catenin complex or by the oncogene FAT4 in Epithelial cells (interactions 62-64). Recently, Dsch1 was found to be a FAT4 ligand between mesenchymal cell contacts [43,113]. Since the ligands are unknown for Epithelial cells, we assumed in this network the activation of FAT4 by a generic cell-cell contact ligand. We also included the reported transcriptional inhibition of FAT4 gene expression by ERK in MCF-10A cell line (interaction 42) [130].

#### 5.1.4. Logical functions

For the high degree of cell-cell adhesion (value 2), we define the logical rule with an “AND” rule between E-cadherin\_mem and b-Cat\_mem at their maximum levels. This was based on the fact that mature and strong adhesion has high E-cadherin concentration in the membrane, bound to  $\beta$ -catenin (E-cadherin/ $\beta$ -catenin complex) [98,99,118,174]. For the intermediate cell-cell adhesion strength (value 1), we assume that the intermediate level of E-cadherin in the membrane (value 1) is enough to ensure this level based on the homophilic binding capacity of E-cadherin [101,175]. This is also supported by reports in *Drosophila* and MDCK cell lines [98–100]. Thus, we set logical rules to allow all remaining combinations of E-Cad\_mem and b-Cat\_mem states except for the ones with E-cad\_mem value of 0 (no homophilic binding). For E-cadherin\_mem, the logical functions were defined accounting for the role of p120 and RAP1 as modulators of E-cadherin concentration at the membrane [101,141,176,177]. Here, we account for the transcriptional activation of E-cadherin (CDH1) as a requisite for both high and intermediate levels [178]. Since p120 stabilizes E-cadherin in the membrane and Rap1 mediates the transport of E-cadherin to the membrane, we assume that high concentration of E-cadherin (value 2) is given by the combination of all activators using “AND” rules to link the regulators [176–178]. The function for the intermediate level of E-cadherin (value 1) was set to describe a particular scenario where E-cadherin gene expression is active and RAP1 is the only mediator of the transport to the membrane [99,100,141]. For the Focal\_Adhesion\_Dynamics, we chose to link all regulatory nodes through “AND” rule because all these regulators participate in the biochemical mechanisms that ensures an high formation and turnover of the focal adhesions [154–158]. Most of the logical functions were based on reported experimental evidence that suggests direct effects (activation or inhibition) not dependent on other regulators. In these cases, we link the regulators through “OR” if its an activation and “AND NOT” if its an inhibitor (see Table S3). We assume that inhibitions are

always dominant over the activations except if there is evidence otherwise. For example, in the functions for b-Cat\_mem and p120 we considered that RPTP activation is dominant over all inhibitors because the activity of RPTPs is about 3 orders of magnitude higher than the activity of RTK [42]. In some cases, the “AND” rule was used between 2 or more activators that conjugate the transcription of a protein and its activation by post-translational mechanisms. The activation of the nodes AKT, NfκB and HIF1α are 3 examples of regulators that contain these types of regulatory mechanisms (see Table S3). For some nodes, we found that they were regulated by co-activation/inhibition biochemical mechanisms. Thus, we used “AND” rules to link activators and “OR NOT” rules for inhibitions for these nodes. ITGab, CDH1, DVL, NfκB, PI3K, RAP1, MEK, JNK and SMAD are examples of nodes that are controlled by co-regulatory mechanisms (see Table S3). The nodes FAK\_SRC, SNAI and TCF\_LEF were also under co-regulatory mechanisms but with a higher degree of complexity. For FAK/SRC complex, the logical function was set to describe the biochemical mechanisms of activation of FAK/SRC such as is able to result in high formation and turnover of focal adhesions. These mechanisms include sequential phosphorylation and dephosphorylation on specific tyrosine residues of FAK [140,179]. The initial step for the activation is through integrin signalling activation in combination with the receptor tyrosine kinases (MET or EGFR), which are necessary to remove FAK auto-inhibition (Y397 and Y194) and promote the binding affinity to SRC [140,161,180], where we write it as: ITGab “AND” (MET “OR” EGFR). Next, SRC activation is necessary to catalyse the phosphorylation of FAK on tyrosine residues Y576, Y577, and Y925, activating the kinase activity of the complex [140,155]. In addition, we only found evidence for SRC activation (phosphorylation at Y418) by either EGFR or MET suggesting that a free SRC activation is required as initial step [94,181]. We assume that SRC effect on the activation of FAK/SRC can be either through the activation by its free form or the complex. Thus, we write a new condition for the activation (SRC “OR” FAK\_SRC) linked by an “AND” with the first condition. Finally, we add another “AND” rule to

ERK based on the reports that showed ERK activation promotes high focal adhesion dynamics through the dephosphorylation of FAK [94,179].

SNAI logical function was set accounting for two alternative mechanisms of expressing snail proteins enough to cause EMT (connected by an “OR” rule). One is the combination of a transient activation of SNAI1 followed by SNAI2 expression [182,183]. In this case, ERK is required for the activation of both SNAI1 and SNAI2 through AP-1 induction [80,184]. Experiments showed that SMAD is a co-activator of SNAI1 expression indicating that it cannot cause an effect alone [184,185]. On the other hand, notch signalling and TCF/LEF induced transcription factors are found to directly bind to SNAI2 promoter and stimulate gene expression [110,186]. Taken all these information together, we wrote the rules for this mechanism as: ERK “AND” SMAD “AND” (CSL “OR” TCF\_LEF). The other mechanism of activation is by a the SMAD induced sustained expression of SNAI1 [184]. This mechanism requires both JNK and NFkB to remove the negative feedback caused by SNAIL1 protein [170,184]. This, together with the co-activation by SMAD, supports the choice of linking JNK, SMAD and Nfkb through “AND” rule. Finally, STAT3, PAK and HIF1a are regulators which are reported to play a role in preventing the repression of SNAI1 by GSK3b phosphorylation [166,187,188]. Thus the term (STAT3 | HIF1a | PAK | !GSK3B) was included in the logical function connected by and “AND” operator to the rules for the 2 mechanisms.

TCF\_LEF function was also set to consider two types of activation mechanisms (connected through “OR”). One mechanism is the translocation of catenins to the nucleus caused by their detachment from the membrane and cytoplasm concentration [189]. This does not require specific phosphorylation on beta-catenin residues. For this mechanism we connect the negation of membrane form of p120 (“NOT” p120) to b-Cat\_nuc through “AND” rule. The other mechanism reflects the constitutive activation of TCF/LEF transcription factor caused by an enhanced accumulation and retention of phosphorylated beta-catenin in the nucleus. This process is promoted by specific phosphorylation of

beta-catenin residues by both MET and AKT [152,153,190]. To describe this mechanism, we assume a self-activation by TCF/LEF and consider that it depends simultaneously on the active MET and AKT by connecting through “AND” rule.

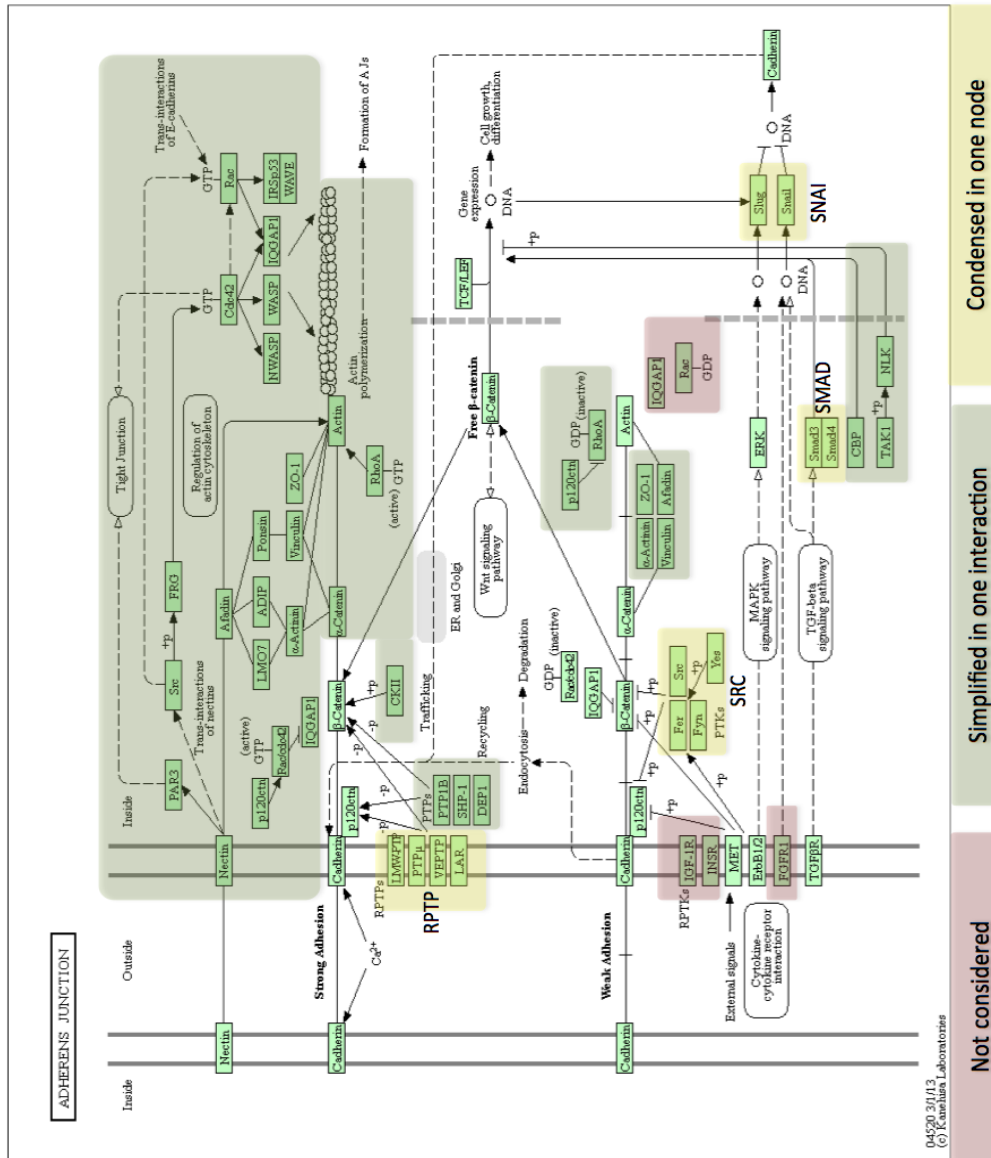
#### **5.1.5. Comparison with a published EMT network model**

Here, we compare our model with the published logical network model of hepatocarcinoma developed by Steinway et al. [31]. This model shares a total of 30 nodes with our model and contains others that we did not include (Table S5). These nodes are the components of the Hedgehog signalling pathway, 3 different growth factor receptors, and additional transcription factors that regulate E-cadherin expression. However, their effects were taken into account in our model by aggregating into one node or considering a direct interaction in order to keep our model smaller (details in Table S5). On the other hand, we included in our model additional nodes and interactions not accounted in the Steinway's model, which are: The Hippo signalling (FAT4, LATS and YAP/TAZ); The integrin signalling (ITGAb, FAK/SRC, JNK, RAP1); The receptor tyrosine phosphatases (RPTP); and regulators of the E-cadherin mediated cell-cell adhesion p120 and CK1. In terms of microenvironment signals, we accounted for the ones in the Steinway model and included additional microenvironments, the two neighbouring cell-cell contact signals (RPTPL and FAT4L), the IL6 and the ECM stiffness (ECM). To further compare the models, we computed the stable states of the Steinway's model (section 2.5.4) and compared the representative stable states with the ones for Epithelial and Mesenchymal adhesion of our model (Figure S9). The comparison of Epithelial stable state (Figure S9A and B) indicate that both models are in agreement, where there is two stable states that are equivalent (see Figure S5). However, the activation of STAT3, Notch and MAPK

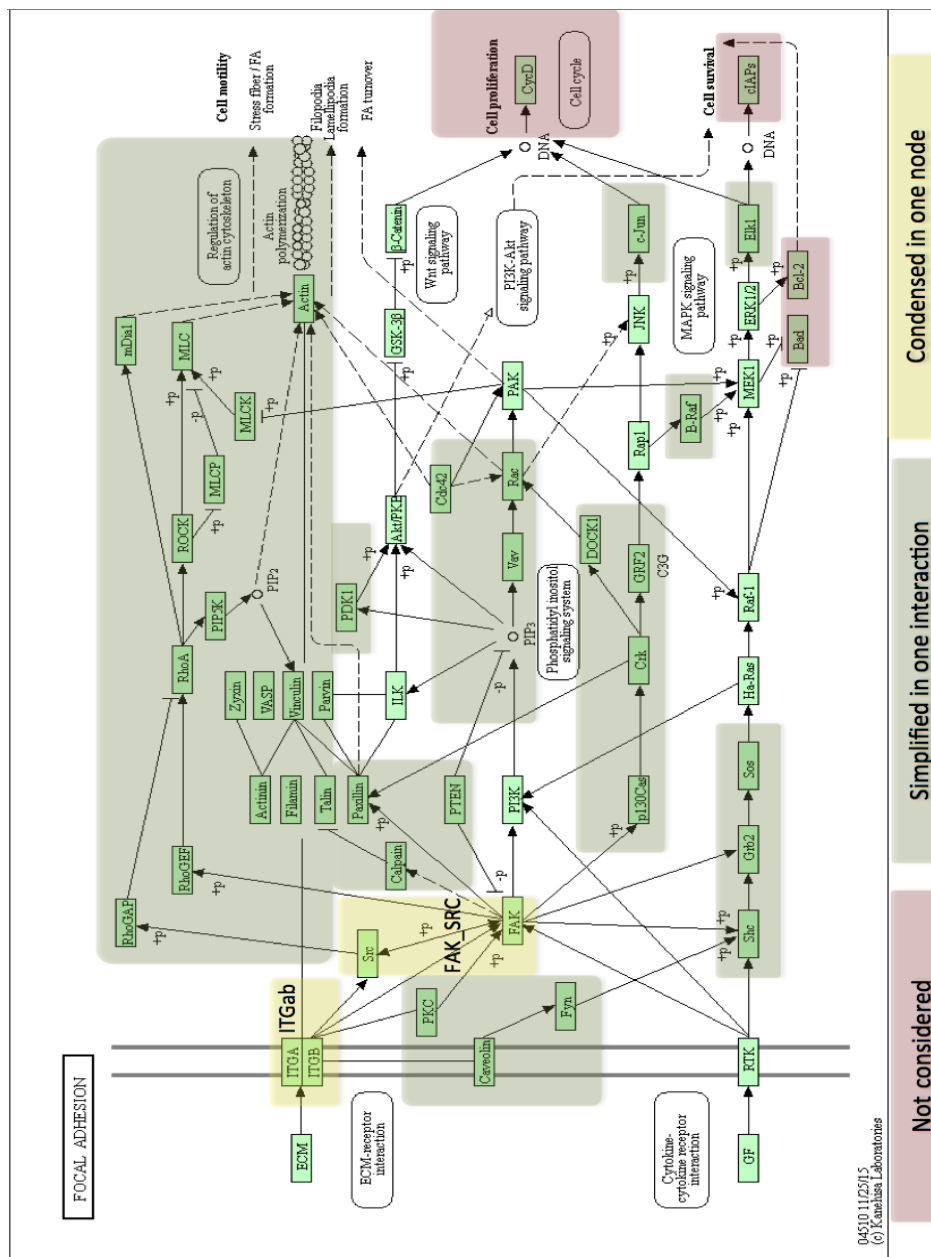
signalling depending on the microenvironment conditions is not obtained in the Steinway model, making this a new prediction. For the Mesenchymal stable states, the activity in both models is also in agreement (Figure S9C and D) , except for the Notch signalling and NFkB activation. In this discrepancy, our model predicts basal activity for these two components and in the Steinway model predicts the opposite. The reason for this difference is because we included in our model the reported inhibition of Notch by DVL, a Wnt signalling component that is activated in all Mesenchymal stable states [191]. For the differences in NFkB activation, this is because we considered the reported NFkB sustained activation due to Notch signalling, whereas Steinway model considered just an initial activation by AKT [192]. We also obtain differences in the activity of SRC in the Mesenchymal stable states because we have considered in our model two different forms of SRC, the free and the complexed with FAK. Thus SRC in Steinway model reflected any form of SRC which is compatible with both SRC or FAK-SRC. Among the regulators not included in the Steinway model, the RPTP and the Hippo signalling components FAT4, LATS and YAP/TAZ, are predicted to have distinct activity status between Epithelial and Mesenchymal adhesion phenotypes. This indicates the need to further analyze these regulators because they could play a role in cell adhesion regulation during EMT.



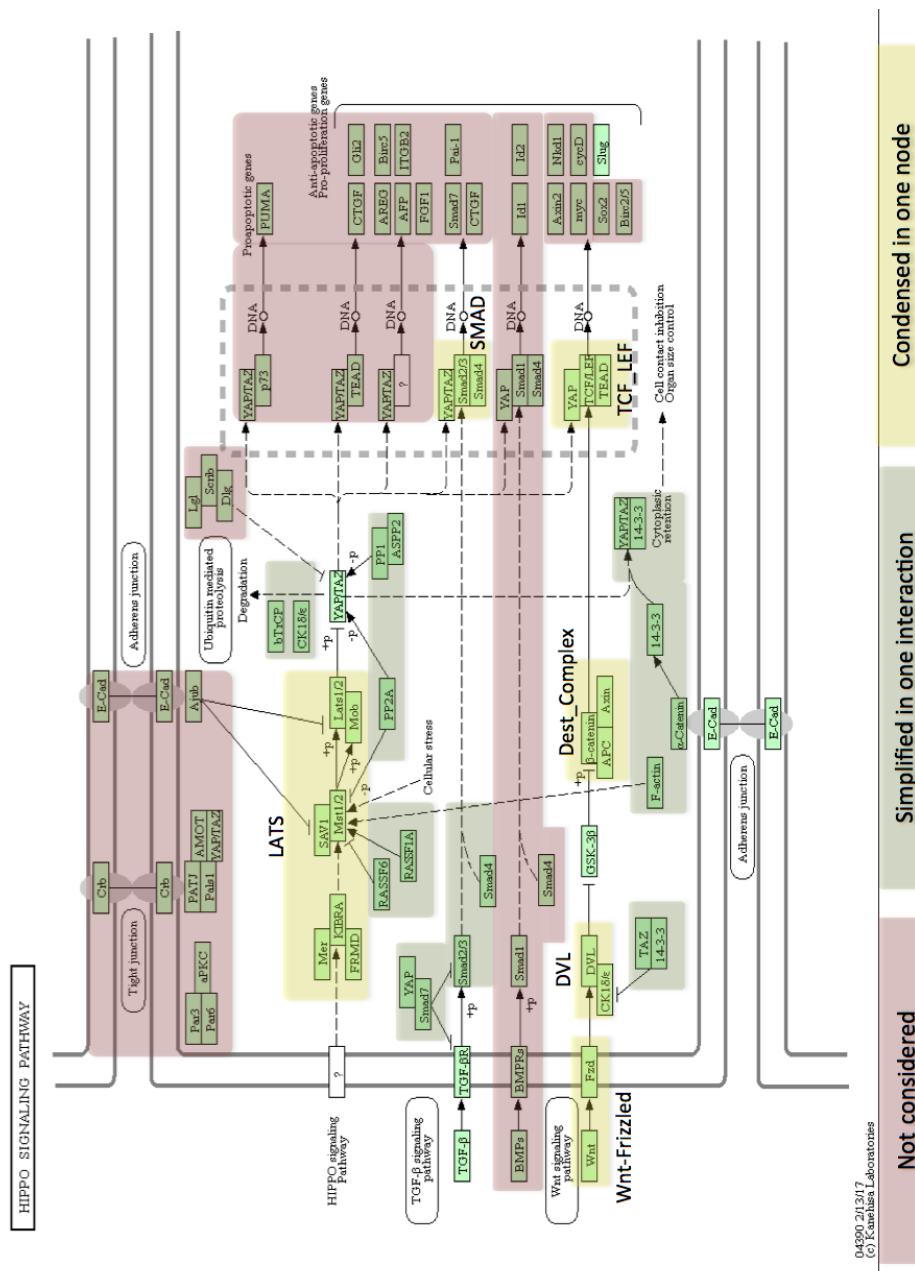
## 5.2. Supplemental figures



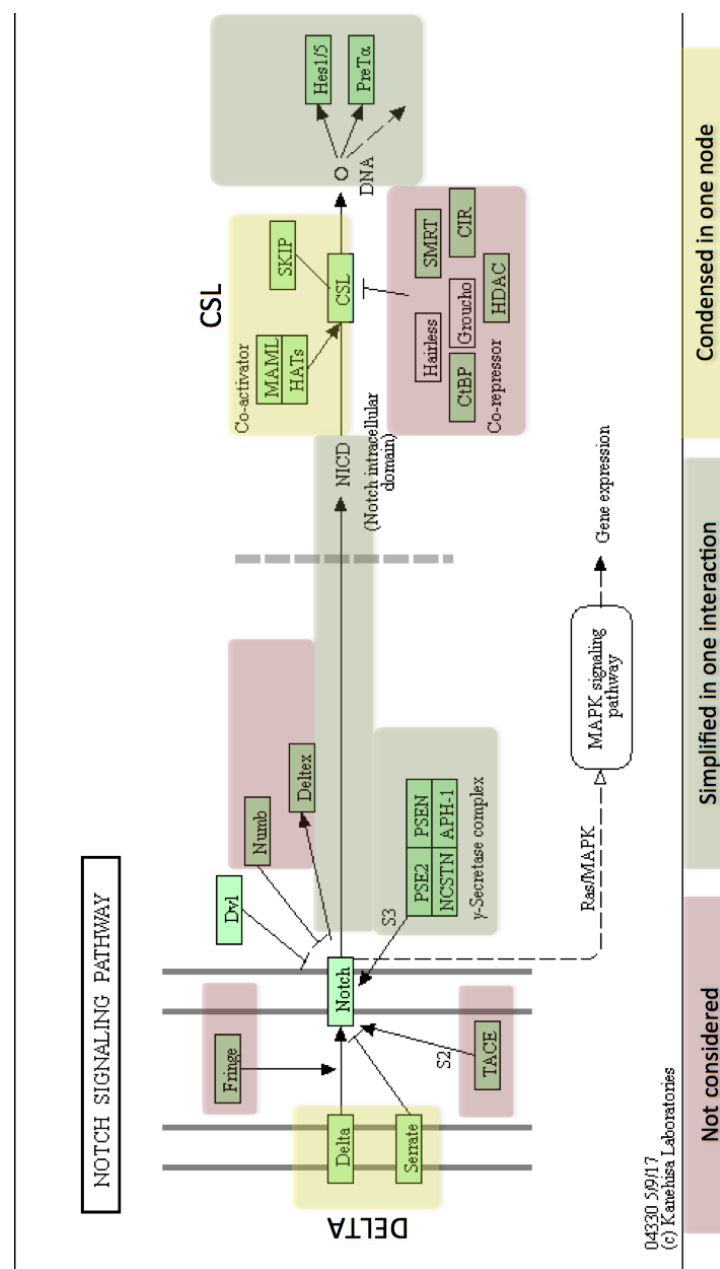
**Figure S1. Adherens junctions regulatory map of KEGG PATHWAYS database (map hsa04520).** Simplifications on the regulatory network presented in section 3.2 are depicted in colours. Nodes names of the network in 3.2 that condensed regulatory components are identified next to the cluster in yellow. Neglected regulatory components are depicted in light red and regulatory mechanisms that were condensed in one interaction in green.



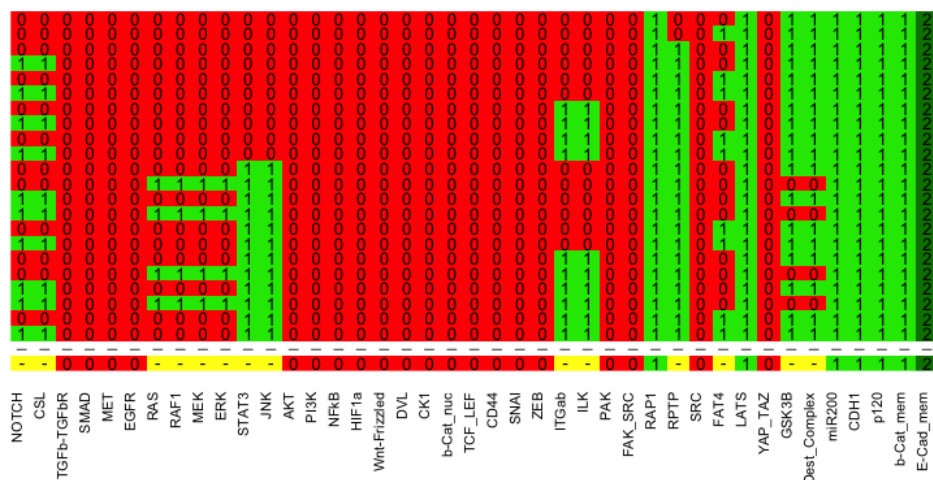
**Figure S2. Focal adhesion regulatory map of KEGG PATHWAYS database (map hsa04510).** Simplifications on the regulatory network presented in section 3.2 are depicted in colours. Nodes names of the network in 3.2 that condensed regulatory components are identified next to the cluster in yellow. Neglected regulatory components are depicted in light red and regulatory mechanisms that were condensed in one interaction in green.



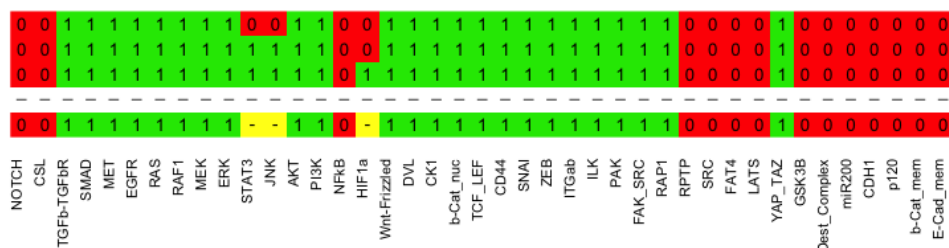
**Figure S3 Hippo signalling map of KEGG PATHWAYS database (map hsa04590).** Simplifications on the regulatory network presented in section 3.2 are depicted in colours. Nodes names of the network in 3.2 that condensed regulatory components are identified next to the cluster in yellow. Neglected regulatory components are depicted in light red and regulatory mechanisms that were condensed in one interaction in green.



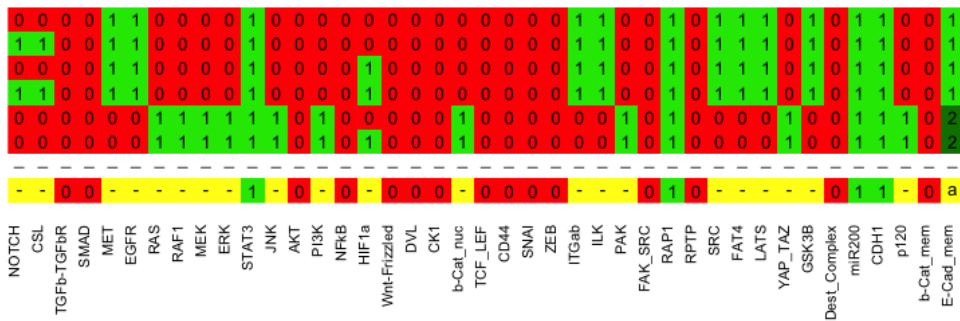
**Figure S4 NOTCH signalling map of KEGG PATHWAYS database (map hsa04530).** Simplifications on the regulatory network presented in section 3.2 are depicted in colours. Nodes names of the network in 3.2 that condensed regulatory components are identified next to the cluster in yellow. Neglected regulatory components are depicted in light red and regulatory mechanisms that were condensed in one interaction in green.



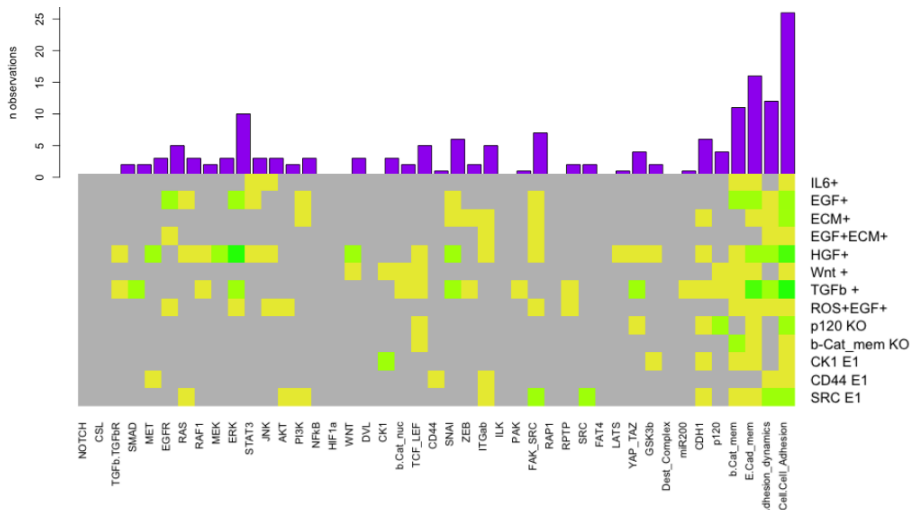
**Figure S5. High cell-cell adhesion and basal focal adhesion dynamics (C=2 & F= 0) stable states.** The stable states were computed as described in section 2.3.2. A representative stable state that indicates the conserved and variable activity of the nodes is shown below dashed line. Red indicates basal activity, green indicate high (Boolean nodes) or intermediate activity (Multi-valued nodes), dark green indicate high activity (Multi-valued nodes). In representative stable state, red and green indicates nodes with conserved activity, and the yellow indicates nodes with variable activity. The symbol (-) in nodes in yellow indicate that all values are identified in the stable states. The letters in yellow indicate exceptions for values 0 (a), 1 (b) and 2 (c).



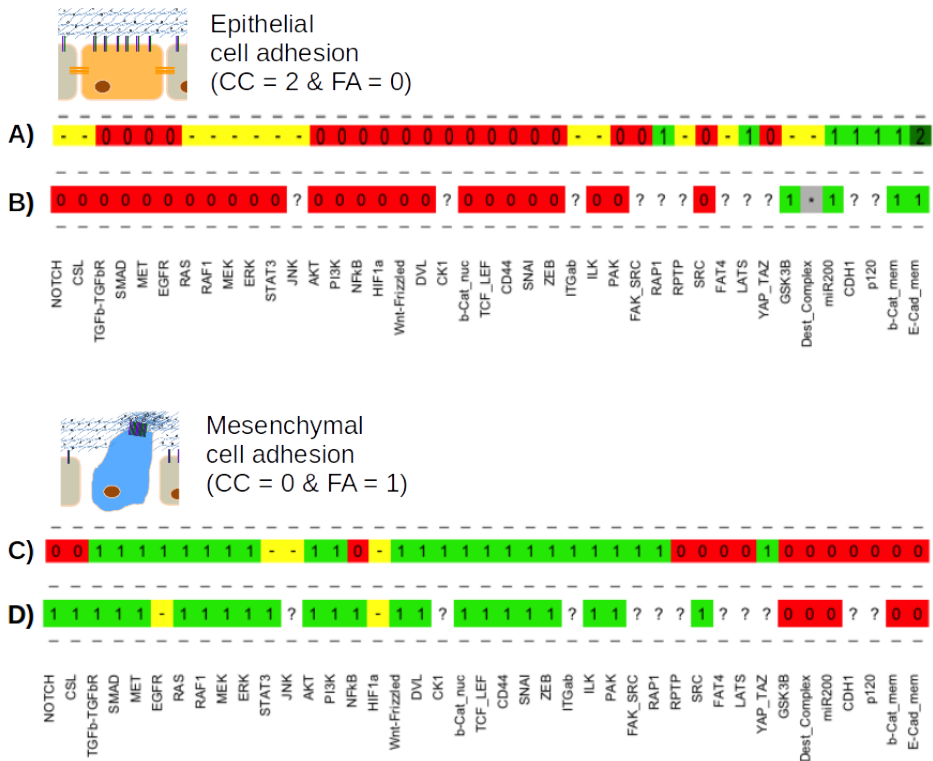
**Figure S6. Basal cell-cell adhesion and high focal adhesion dynamics (C=2 & F= 0) stable states.** The stable states were computed as described in section 2.3.2. A representative stable state that indicates the conserved and variable activity of the nodes is shown below dashed line. Red indicates basal activity, green indicate high (Boolean nodes) or intermediate activity (Multi-valued nodes), dark green indicate high activity (Multi-valued nodes). In representative stable state, red and green indicates nodes with conserved activity, and the yellow indicates nodes with variable activity. The symbol (-) in nodes in yellow indicate that all values are identified in the stable states. The letters in yellow indicate exceptions for values 0 (a), 1 (b) and 2 (c).



**Figure S7. Intermediate cell-cell adhesion and basal focal adhesion dynamics ( $C=1$  &  $F=0$ ) stable states.** The stable states were computed as described in section 2.3.2. A representative stable state that indicates the conserved and variable activity of the nodes is shown below dashed line. Red indicates basal activity, green indicate high (Boolean nodes) or intermediate activity (Multi-valued nodes), dark green indicate high activity (Multi-valued nodes). In representative stable state, red and green indicates nodes with conserved activity, and the yellow indicates nodes with variable activity. The symbol (-) in nodes in yellow indicate that all values are identified in the stable states. The letters in yellow indicate exceptions for values 0 (a), 1 (b) and 2 (c).



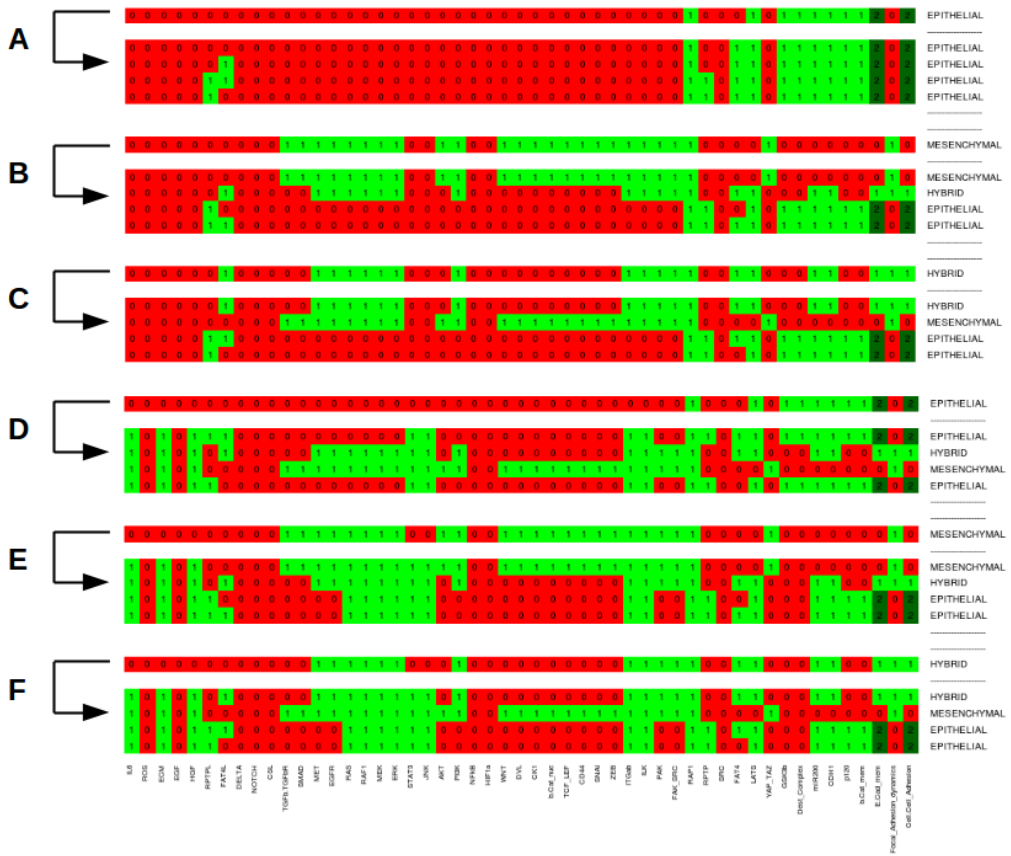
**Figure S8. Number of observations in the experiments on epithelial cell lines used for model validation.** The values correspond to the experiments in Table 5. Yellow indicate 1 collected observation, light green 2 and green 3 independent observations from different references. The purple bars indicate the total observations for each node.



**Figure S9. Comparison of model stable states for Epithelial and Mesenchymal phenotypes with the ones of a published EMT model.** Epithelial and Mesenchymal stable states of our model indicated in A and C, respectively. Epithelial and Mesenchymal stable states of the model of Steinway [31] indicated in B and D, respectively. The representative stable states were computed as described in section 2.3.2. Red indicate basal activity, green indicate high (Boolean nodes) or intermediate activity (Multi-valued nodes), dark green indicate high activity (Multi-valued nodes). Yellow indicates nodes with variable activity, grey any value (\*), and the remaining colours conserved activity. The symbol (-) in nodes in yellow indicate that all values are identified in the stable states. The letters in yellow indicate exceptions for values 0 (a), 1 (b) and 2 (c). The symbol ? in white indicates nodes that are only included in our model.







**Figure S12. Simulations of the model with FAT4 perturbation (FAT4[ERK@0]) for normal A-C and chronic inflammation conditions D-F. Initial and final states of simulations are indicated by the arrows.**

### 5.3. Supplemental tables

**Table S1. Network model components.**

<b>Node</b>	<b>Full name</b>	<b>Regulatory Module</b>	<b>Uniprot NCBI*</b>
<b>AKT</b>	RAC-alpha serine/threonine-protein kinase 1/2	AKT signaling	P31749 P31751
<b>b-Cat_mem</b>	Unphosphorylated form of $\beta$ -catenin in the membrane	Adherens junction	P35222
<b>b-Cat_nuc</b>	$\beta$ -catenin in the nucleus	Wnt signalling	P35222
<b>CD44</b>	CD44 antigen ligand	Focal adhesion	P16070
<b>CDH1</b>	E-cadherin gene	Adherens junction	999*
<b>CK1</b>	Casein kinase 1	Wnt signalling	P48729
<b>CSL</b>	C protein binding factor 1/Suppressor of Hairless/Lag-1 complex.	Notch signalling	[91]
<b>DELTA</b>	Delta-like protein ligand	Notch signalling	P78504 O00548
<b>Dest_Complex</b>	$\beta$ -catenin destruction complex	Wnt Signalling	[193]
<b>DVL</b>	Human homologue of segment polarity protein dishevelled 1	Wnt signalling	O14640
<b>E-Cad_mem</b>	E-cadherin in the membrane	Adherens junction	P12830
<b>ECM</b>	Collagen/fibronectin induced extracellular matrix stiffness	Focal adhesion	[194]
<b>EGF</b>	Epidermal growth factor	Adherens junction	[149]
<b>EGFR</b>	Human epidermal growth factor receptor family members 1 and 2	Adherens junction	P00533 P04626
<b>ERK</b>	Mitogen-activated kinase	MAPK signalling	P28482
<b>FAK_SRC</b>	Focal adhesion kinase (FAK) and tyrosine-protein kinase Src (SRC) complex activity dynamics	Focal adhesion	[140]
<b>FAT4</b>	Human homologue of Drosophila FAT4 protein receptor	Hippo signalling	Q6V0I7
<b>FAT4L</b>	Generic Dsch1-like FAT4 ligand proteins	Hippo signalling	[41,43]
<b>GSK3B</b>	Glycogen synthase kinase 3 beta	Wnt Signalling	P49841
<b>HGF</b>	Hepatocyte growth factor	Adherens junction	P14210
<b>HIF1a</b>	Hypoxia-inducible factor 1-alpha	HIF1 signalling	Q16665
<b>IL6</b>	Interleukin-6	JAK-STAT signalling	P05231
<b>ILK</b>	Integrin-linked protein kinase	Focal adhesion	Q13418
<b>ITGab</b>	Transmembrane integrins alpha5beta1	Focal adhesion	[154,195]

<b>JNK</b>	c-jun N-terminal kinase	Focal adhesion	{Formatting Citation}
<b>LATS</b>	Large tumour suppressor complex LATS1/2	Hippo signalling	O95835 Q9NRM7
<b>MEK</b>	Dual specificity mitogen-activated kinase kinase	MAPK signalling	Q02750
<b>MET</b>	Hepatocyte growth factor receptor	Adherens junction	P08581
<b>miR200</b>	Micro RNA 200 family	Adherens junction	[196]
<b>NFkB</b>	Nuclear factor NF-kappa-B complex	AKT signaling	Q04206 P19838 Q00653
<b>NOTCH</b>	Neurogenic locus notch homolog protein	Notch signalling	P46531
<b>p120</b>	Unphosphorylated form of $\delta$ -catenin in the membrane	Adherens junction	O60716
<b>PAK</b>	Serine/threonine-protein kinase PAK1	Focal adhesion	Q13153
<b>PI3K</b>	Phosphatidylinositol 4,5-bisphosphate 3-kinase	AKT signaling	P42336
<b>RAF1</b>	Proto-oncogene serine/threonine-protein kinase Raf 1	MAPK signalling	P04049
<b>RAP1</b>	Ras-related protein Rap-1	Focal adhesion	P62834
<b>RAS</b>	Ras GTPase family of proteins	MAPK signalling	P01116 P01112
<b>ROS</b>	Reactive Oxygen chemical Species	HIF1 signalling	[22,164]
<b>RPTP</b>	Receptor-type Protein-Tyrosine Phosphatase family members	Adherens junction	Q12913 Q15262 P28827
<b>RPTPL</b>	Generic Receptor Protein-Tyrosine Phosphatase Ligand	Adherens junction	[44]
<b>SMAD</b>	SMAD 2, 3 and 4 signal transducer complex	TGF- $\beta$ signalling	P84022 Q13485 Q15796
<b>SNAI</b>	Human homolog Zinc finger protein snail 1 and 2	Adherens junction	O95863 O43623
<b>SRC</b>	Free form of proto-oncogene tyrosine-protein kinase Src	Adherens junction	P12931
<b>STAT3</b>	Signal transducer and activator of transcription 3	JAK-STAT signalling	P40763
<b>TCF_LEF</b>	Transcription factor TCF/LEF family	Wnt signalling	IPR024940 Q9NQB0 Q9UJU2
<b>TGFb-TGFbR</b>	Transforming growth factor beta 1 ligand/receptor complex	TGF- $\beta$ signalling	P01137 P36897

<b>Wnt-Frizzled</b>	Human homolog Wnt and Frizzled ligand/receptor complex	Wnt signalling	Q9UP38 P04628
<b>YAP_TAZ</b>	Transcriptional co-activator YAP/TAZ	Hippo signalling	P46937 Q9GZV5
<b>ZEB</b>	Zinc finger E-box-binding homeobox protein 1/2	Adherens junction	P37275 O60315

**Table S2. Interactions of the network and their effects.** N denotes the reference number for each interaction in the model (section 3.2). Interactions that are present in KEGG PATHWAYS are indicated with the letter K, the ones identified in the Atlas of Cancer Signalling Network (ACSN) with the letter A, and the ones included in the EMT network model of Steinway [31] with the letter S. References on each interaction are in model\_documentation.xhtml file available in gitub (see section 2.2.5 to get file).

<b>N</b>	<b>Target</b>	<b>Regulator</b>	<b>Effect</b>	<b>Slowest Mechanism</b>	<b>Main source</b>
1	AKT	PI3K	Activation	Phosphorylation	K, S
2	AKT	SNAI	Activation	Transcription	A
3	AKT	ILK	Activation	Phosphorylation	K, S
4	b-Cat_mem	RPTP	Activation	Dephosphorylation	K
5	b-Cat_mem	CDH1	Activation	Molecular binding	K, S
6	b-Cat_mem	JNK	Inhibition	Phosphorylation	L
7	b-Cat_mem	MET	Inhibition	Phosphorylation	K
8	b-Cat_mem	SRC	Inhibition	Phosphorylation	K
9	b-Cat_mem	FAK_SRC	Inhibition	Phosphorylation	K
10	b-Cat_nuc	YAP_TAZ	Activation	Transport	K
11	b-Cat_nuc	Dest_Complex	Inhibition	Transport	K, S
12	b-Cat_nuc	b-Cat_mem	Inhibition	Transport	K, S
13	CD44	TCF_LEF	Activation	Transcription	S
14	CDH1	SNAI	Inhibition	Transcription	A, S
15	CDH1	ZEB	Inhibition	Transcription	A, S
16	Cell-Cell Adhesion	E-Cad_mem	Activation	Molecular binding	K
17	Cell-Cell Adhesion	b-Cat_mem	Activation	Molecular binding	K
18	CK1	Wnt-Frizzled	Activation	Molecular binding	K
19	CK1	Dest_Complex	Inhibition	Molecular binding	K
20	CSL	NOTCH	Activation	Transcription	K, S
21	Dest_Complex	b-Cat_mem	Activation	Molecular binding	K, S

22	Dest_Complex	GSK3B	Activation	Phosphorylation	K, S
23	DVL	YAP_TAZ	Activation	Molecular binding	K
24	DVL	Wnt-Frizzled	Activation	Phosphorylation	K, S
25	E-Cad_mem	RAP1	Activation	Transport	A
26	E-Cad_mem	p120	Activation	Molecular binding	K
27	E-Cad_mem	CDH1	Activation	Transport	S
28	EGFR	CSL	Activation	Transcription	L
29	EGFR	ITGab	Activation	Molecular binding	L
30	EGFR	EGF	Activation	Molecular binding	K, S
31	EGFR	TGFb-TGFbR	Activation	Phosphorylation	L
32	EGFR	SRC	Activation	Phosphorylation	S
33	EGFR	RPTP	Inhibition	Dephosphorylation	L
34	ERK	MEK	Activation	Phosphorylation	K, S
35	FAK_SRC	ERK	Activation	Dephosphorylation	L
36	FAK_SRC	ITGab	Activation	Molecular binding	K
37	FAK_SRC	EGFR	Activation	Phosphorylation	L
38	FAK_SRC	MET	Activation	Phosphorylation	K
39	FAK_SRC	SRC	Activation	Phosphorylation	K
40	FAK_SRC	FAK_SRC	Activation	Phosphorylation	K
41	FAT4	FAT4L	Activation	Molecular binding	L
42	FAT4	ERK	Inhibition	Transcription	A
43	Focal Adhesion Dynamics	PAK	Activation	Phosphorylation	K
44	Focal Adhesion Dynamics	ITGab	Activation	Molecular binding	K
45	Focal Adhesion Dynamics	FAK_SRC	Activation	Phosphorylation	K
46	GSK3B	DVL	Inhibition	Phosphorylation	K, S
47	GSK3B	ERK	Inhibition	Phosphorylation	A, S
48	GSK3B	AKT	Inhibition	Phosphorylation	K, S
49	HIF1a	STAT3	Activation	Transcription	K
50	HIF1a	ROS	Activation	Oxidation	K
51	HIF1a	NFkB	Activation	Transcription	K
52	ILK	ITGab	Activation	Molecular binding	K
53	ITGab	RAP1	Activation	Transport	A
54	ITGab	CD44	Activation	Molecular binding	S

55	ITGab	TGFb-TGFbR	Activation	Molecular binding	L
56	ITGab	ZEB	Activation	Transcription	A
57	ITGab	ECM	Activation	Molecular binding	K
58	JNK	TGFb-TGFbR	Activation	Phosphorylation	L
59	JNK	RAS	Activation	Phosphorylation	K
60	JNK	IL6	Activation	Phosphorylation	K
61	JNK	ERK	Inhibition	Phosphorylation	L
62	LATS	FAT4	Activation	Molecular binding	K
63	LATS	E-Cad_mem	Activation	Molecular binding	K
64	LATS	b-Cat_mem	Activation	Molecular binding	K
65	LATS	PI3K	Inhibition	Phosphorylation	K
66	MEK	PAK	Activates	Phosphorylation	K, S
67	MEK	RAP1	Activation	Phosphorylation	K
68	MET	CD44	Activation	Phosphorylation	S
69	MET	ITGab	Activation	Phosphorylation	L
70	MET	HGF	Activation	Molecular binding	K, S
71	MET	RPTP	Inhibition	Dephosphorylation	L
72	miR200	SNAI	Inhibition	Transcription	A, S
73	miR200	ZEB	Inhibition	Transcription	A, S
74	NFkB	CSL	Activation	Transcription	A
75	NFkB	AKT	Activation	Phosphorylation	K, S
76	NFkB	ROS	Activation	Oxidation	K
77	NFkB	SMAD	Activation	Transcription	A
78	NOTCH	DELTA	Activation	Molecular binding	K, S
79	NOTCH	DVL	Inhibition	Molecular binding	K
80	p120	RPTP	Activation	Dephosphorylation	K
81	p120	CK1	Inhibition	Phosphorylation	L
82	p120	MET	Inhibition	Phosphorylation	K
83	p120	SRC	Inhibition	Phosphorylation	K
84	PAK	MEK	Activation	Phosphorylation	K
85	PAK	PI3K	Activation	Phosphorylation	K
86	PAK	FAK_SRC	Activation	Phosphorylation	K
87	PI3K	RAS	Activation	Phosphorylation	K , S
88	PI3K	IL6	Activation	Phosphorylation	K
89	PI3K	FAK_SRC	Activation	Phosphorylation	K

90	PI3K	E-Cad_mem	Inhibition	Molecular binding	L
91	PI3K	b-Cat_mem	Inhibition	Molecular binding	L
92	RAF1	RAS	Activation	Phosphorylation	K, S
93	RAF1	MEK	Activation	Phosphorylation	K, S
94	RAF1	ERK	Inhibition	Phosphorylation	A, S
95	RAP1	CDH1	Activation	Transport	L
96	RAP1	FAK_SRC	Activation	Phosphorylation	K
97	RAS	EGFR	Activation	Phosphorylation	K, S
98	RAS	TGFb-TGFbR	Activation	Phosphorylation	K, S
99	RAS	MET	Activation	Phosphorylation	K, S
100	RAS	SRC	Activation	Phosphorylation	K, S
101	RAS	IL6	Activation	Phosphorylation	K, S
102	RAS	FAK_SRC	Activation	Phosphorylation	K
103	RAS	GSK3B	Inhibition	Phosphorylation	K, S
104	RPTP	RPTPL	Activation	Molecular binding	L
105	RPTP	ROS	Inhibition	Oxidation	L
106	RPTP	SRC	Inhibition	Transcription	L
107	SMAD	YAP_TAZ	Activation	Molecular binding	K
108	SMAD	TGFb-TGFbR	Activation	Phosphorylation	K, S
109	SMAD	ILK	Activation	Phosphorylation	A
110	SNAI	ERK	Activation	Transcription	A, S
111	SNAI	CSL	Activation	Transcription	A, S
112	SNAI	JNK	Activation	Phosphorylation	A, S
113	SNAI	PAK	Activation	Transcription	A, S
114	SNAI	STAT3	Activation	Transcription	A, S
115	SNAI	HIF1a	Activation	Transcription	A, S
116	SNAI	SMAD	Activation	Transcription	A, S
117	SNAI	TCF_LEF	Activation	Transcription	A, S
118	SNAI	NFkB	Activation	Transcription	A, S
119	SNAI	GSK3B	Inhibition	Phosphorylation	A, S
120	SRC	EGFR	Activation	Phosphorylation	A, S
121	SRC	MET	Activation	Phosphorylation	K
122	SRC	FAK_SRC	Inhibition	Molecular binding	L
123	STAT3	JNK	Activation	Phosphorylation	A
124	STAT3	SRC	Activation	Phosphorylation	K, S

125	TCF_LEF	b-Cat_nuc	Activation	Molecular binding	K, S
126	TCF_LEF	AKT	Activation	Phosphorylation	L
127	TCF_LEF	MET	Activation	Phosphorylation	L
128	TCF_LEF	p120	Inhibition	Transcription	L
129	TCF_LEF	TCF_LEF	Activation	Molecular binding	L
130	TGFb-TGFbR	TCF_LEF	Activation	Transcription	K, S
131	Wnt-Frizzled	TCF_LEF	Activation	Transcription	K, S
132	YAP_TAZ	LATS	Inhibition	Phosphorylation	K
133	ZEB	SMAD	Activation	Transcription	A, S
134	ZEB	miR200	Inhibition	Transcription	A, S

**Table S3. Model logical functions and the correspondent type of regulatory mechanisms (R.M.).** The Logical operators are represented by the symbols & (AND), | (OR) and ! (NOT). Post-translational mechanisms are indicated with PT and transcriptional mechanisms with T.

Node	Logical function	R.M.
<b>NOTCH = 1</b>	DELTA & ! DVL	PT
<b>CSL =1</b>	NOTCH	T + PT
<b>TGFb-TGFbR =1</b>	TCF_LEF	T
<b>SMAD = 1</b>	TGFb-TGFbR & ILK & YAP_TAZ	PT
<b>MET = 1</b>	(CD44   HGF   ITGab) & !RPTP	PT
<b>EGFR = 1</b>	(EGF   CSL   TGFb-TGFbR   SRC   ITGab) & !RPTP	PT
<b>RAS = 1</b>	(TGFb-TGFbR   MET   FAK_SRC   IL6   SRC   EGFR) & !GSK3B	PT
<b>RAF1 = 1</b>	RAS & !ERK	PT
<b>MEK = 1</b>	(RAF1 & RAP1)   PAK	PT
<b>ERK = 1</b>	MEK	PT
<b>STAT3 = 1</b>	JNK   SRC	PT
<b>JNK = 1</b>	(RAS & TGFb-TGFbR & !ERK)   IL6	PT
<b>AKT = 1</b>	(PI3K   ILK) & SNAI	T + PT
<b>PI3K = 1</b>	(RAS   FAK) & !b-Cat_mem & !E-Cad_mem	PT
<b>NfκB = 1</b>	(AKT   ROS) & CSL & SMAD	T+ PT
<b>HIF1a = 1</b>	ROS & (STAT3   NFκB)	T+ PT
<b>Wnt-Frizzled = 1</b>	TCF_LEF	T+ PT



<b>DVL = 1</b>	Wnt-Frizzled & YAP_TAZ	PT
<b>CK1 = 1</b>	Wnt-Frizzled & !Dest_Complex	PT
<b>b-Cat_nuc = 1</b>	!Dest_Complex & YAP_TAZ & !b-Cat_mem	PT
<b>TCF_LEF = 1</b>	(!p120 & b-Cat_nuc)   (TCF_LEF & MET & AKT)	T+ PT
<b>CD44 = 1</b>	TCF_LEF	T
<b>SNAI = 1</b>	(ERK & SMAD & (TCF_LEF   CSL) & !GSK3B)   (JNK & SMAD & NFkB & (STAT3   PAK   HIF1a))	T+ PT
<b>ZEB = 1</b>	!miR200 & SMAD	T+ PT
<b>ITGab = 1</b>	ECM   ((TGFb-TGFbR   CD44   ZEB) & RAP1)	PT
<b>ILK = 1</b>	ITGab	PT
<b>PAK = 1</b>	FAK_SRC   PI3K	PT
<b>FAK_SRC</b>	ITGab & EGFR & MET & ERK & (SRC   FAK_SRC)	PT
<b>RAP1 = 1</b>	FAK_SRC   CDH1	PT
<b>RPTP = 1</b>	RPTPL & !ROS & !SRC	T+ PT
<b>SRC = 1</b>	(MET   EGFR) & !FAK_SRC	PT
<b>FAT4 = 1</b>	FAT4L & !ERK	T+ PT
<b>LATS = 1</b>	(b-Cat_mem:1 & E-Cad_mem & !PI3K)   FAT4	PT
<b>YAP_TAZ = 1</b>	!LATS	PT
<b>GSK3B = 1</b>	!DVL & !ERK & !AKT	PT
<b>Dest_Complex = 1</b>	GSK3B & b-Cat_mem	PT
<b>miR200 = 1</b>	!ZEB & !SNAI	T
<b>CDH1 = 1</b>	!ZEB   !SNAI	T
<b>p120 = 1</b>	(!MET & !SRC & !CK1)   RPTP	PT
<b>b-Cat_mem = 1</b>	(!MET & !SRC & !FAK & !JNK & CDH1)   (RPTP & CDH1)	PT
<b>E-Cad_mem = 2</b>	CDH1 & p120 & RAP1	T+ PT
<b>E-Cad_mem = 1</b>	CDH1 & RAP1 & !p120	
<b>Focal Adhesion Dynamics = 1</b>	ITGab & PAK & FAK_SRC	PT
<b>Cell-cell Adhesion:2</b>	b-Cat_mem & E-Cad_mem:2	PT
<b>Cell-cell Adhesion:1</b>	E-Cad_mem:1   (E-Cad_mem:2 & !b-Cat_mem)	

**Table S4. Parameters used in the simulations of experiments on epithelial cell lines.** Only the nodes with values of high/intermediate activation are presented. The remaining nodes are set to value 0.

Experiments	Model inputs ( $\neq 0$ )	Fixed perturbations	Initial condition ( $\neq 0$ )
IL6+	IL6 = 1	None	Cell-Cell_Adhesion=2 E-Cad = 2 CDH1=1 b-Cat_mem=1 p120=1 miR200=1 GSK3b=1 LATS=1
EGF+	EGF = 1	None	
ECM+	ECM = 1	None	
EGF+ECM+	ECM = 1, EGF = 1	None	
HGF+	HGF = 1	None	
WNT+	None	Wnt-Frizzled = 1	
TGFb+	None	TGFb-TGFbR = 1	
EGF+ROS+	ROS = 1 , EGF = 1	None	
P120 KO	None	p120 = 0	
B-Cat_mem KO	None	b-Cat_mem = 0	
CK1 E1	None	CK1 = 1	
CD44 E1	None	CD44 = 1	
SRC E1	None	SRC = 1	

**Table S5. Nodes of the Steinway EMT logical network model not included in the logical network model of the regulation of cell adhesion properties.**

Nodes	Role in the regulation of EMT	Cell adhesion regulatory network model.
<b>SHH</b> <b>PTCH</b> <b>SMO</b> <b>FUS</b> <b>SUFU</b> <b>GLI</b>	Hedgehog signalling pathway. Demonstrated in Steinway's work to directly joint Wnt signalling with TGF-beta signalling through TCF/LEF activation.	Considered effect and simplified by a direct arc between TCF/LEF and TGF-beta (TCF_LEF $\rightarrow$ TGFb).
<b>PDGFR</b> <b>IGFR</b> <b>FGFR</b>	Activation of MAPK signalling cascade and none receptor tyrosine kinases by growth factors.	Not included in the model. These growth factors share similar signalling cascade activation. The regulatory effect is represented by EGF induced EGFR activation (the most abundant growth factor in epithelial cells).
<b>EGR1</b> <b>c-fos</b>	Transcription factors that mediate SNAI1 activation via ERK.	Considered effect and simplified by a direct arc between ERK and SNAI.
<b>Csn</b> <b>Ikka</b> <b><math>\beta</math>TrCP</b>	Signalling components that mediate NFkB activation by AKT and SNAI1 activation by Nfkb.	Simplified by removing mediators and linking AKT to Nfkb and Nfkb to SNAI with arcs (AKT $\rightarrow$ NFkB $\rightarrow$ SNAI).
<b>SNAI2</b>	Transcription factors that participate in the	Simplified by only considering the two

<b>ZEB2 TWIST1 FOXC2</b>	transcription repression of E-cadherin together with SNAI1 and ZEB1.	main transcription regulators ZEB and SNAI family (two nodes).
<b>NOTCHic HEY1</b>	Participates in the down-regulation of E-cadherin expression through the pathway NOTCH → NOTCHic → CSL → HEY.	The effect is considered via activation of SNAI. Simplified by considering a minimal path (NOTCH → CSL → SNAI).
<b>LIV1</b>	Participates in the activation of SNAI1 via STAT activation.	Considered effect and simplified by a direct arc between STAT3 and SNAI.
<b>LOXL23</b>	Participates in the activation of SNAI1 via HIF1a activation.	Considered effect and simplified by a direct arc between HIF1a and SNAI.
<b>Frizzled TGFbR</b>	Membrane receptors of Wnt and TGF ligands.	Simplified by considering one node for both ligand and receptor (active receptor/ligand complex).
<b>Jagged</b>	Alternative NOTCH activating ligand.	Simplified by considering one node that represents all DELTA-like ligands.
<b>Goosecoid</b>	Transcription factor involved in cell differentiation. Putative auto-regulation mechanism. Involved in EMT by inducing SNAI, ZEB and other TFs.	Not considered in the model because it is not expressed in epithelial cells. Known to be expressed only during development.
<b>CHD1L</b>	DNA damage response signal protein that triggers EMT.	DNA damage response was not considered in the cell adhesion model.
<b>CDC42</b>	Small GTPase that mediates PAK1 activity.	Considered effect and simplified by a direct arc between FAK and PAK.
<b>SOS Grb2</b>	Adapter proteins involved in MAPK signaling propagation.	Considered effect and simplified by a direct arc between growth factor receptor and RAS.
<b>Axin2</b>	Part of the deconstruction complex and considered to be auto-regulated	Simplified by assuming that is expressed and a part of the destruction complex.

**Table S6. Model input conditions (microenvironment conditions) compatible with the model adhesion phenotypes for each combination of cell-cell contact signals (RPTP and FAT4 ligands).** Focal adhesion dynamics is represented by the letter F and cell-cell adhesion by the letter C. The percentage of the remaining microenvironments for a given combination of cell-cell contact signals is indicated in the last column.

Adhesion Phenotypes	Cell-cell contact signals (RPTPL, FAT4L)	Remaining microenvironment conditions (IL6, ROS, ECM, EGF, HGF, DELTA)	%
<b>C=2 &amp; F=0</b>	RPTPL=0 & FAT4L=0 RPTPL=0 & FAT4L=1 RPTPL=1 & FAT4L=0 RPTPL=1 & FAT4L=1	IL6=ECM=EGF=HGF=DELTA=0 IL6=ECM=EGF=HGF=DELTA=0 ROS=0   IL6=ECM=EGF=HGF=DELTA=0 ROS=0   IL6=ECM=EGF=HGF=DELTA=0	3,1 3,1 51,6 51,6
<b>C=1 &amp; F=0</b>	RPTPL=0 & FAT4L=0 RPTPL=0 & FAT4L=1 RPTPL=1 & FAT4L=0 RPTPL=1 & FAT4L=1	IL6=1 & ECM=EGF=HGF=DELTA=0 IL6=1 & ECM=EGF=HGF=DELTA=0   IL6=0 Not stable IL6=ROS=1 & ECM=EGF=HGF=DELTA=0   IL6=0	3,1 53,1 ..... 51,6
<b>C=0 &amp; F=1</b>	RPTPL=0 & FAT4L=0 RPTPL=0 & FAT4L=1 RPTPL=1 & FAT4L=0 RPTPL=1 & FAT4L=1	All combinations All combinations ROS=1 ROS=1	100,0 100,0 50,0 50,0

**Table S7. Model input configurations compatible with physiological scenarios used in simulations.**

Physiological scenarios	Input configuration
<b>Normal tissue</b> Basal level of growth factors, cytokines, ROS and ECM stiffness (low level of collagen type I and fibronectin) [22,23,197].	IL6=ROS=ECM=EGF=HGF=0
<b>Tissue growth</b> High secretion of growth factors by Fibroblasts and adjacent Epithelial cells [149,150,197] Stimulates cell proliferation to balance cell death. Occurs during tissue repair and tissue size homeostasis.	HGF=EGF=1 & IL6=ROS=ECM=0
<b>Chronic Inflammation</b> Accumulation of collagen I, IL6 cytokine and HGF secreted by recruited Fibroblasts and Macrophages to inflammatory site [22,150].	IL6=ECM=HGF=1 & EGF=ROS=0
<b>Hypoxia + chronic inflammation</b> Low levels of oxygen result in high levels of ROS [164]. These conditions together with chronic inflammation are frequently observed in invasive tumours [16,22,23].	IL6=ROS=ECM=HGF=1 & EGF=0

**Table S8. Comparison of the effects of single perturbations on model components with the expression of these proteins in human carcinomas.** The gains of a phenotype in comparison to the unperturbed model phenotypes is identified by the letter **G** and the loss of a phenotype by the letter **L**. Effects that favours or counteract tumour invasion are indicated in red or green, respectively. Perturbations that maintain phenotypes are indicated by the letter **M** and the ones that are not able to cause a gain are indicated by the letter **N**. Focal adhesion dynamics is abbreviated by **F** and cell-cell adhesion by **C**. Protein expression levels of carcinomas in comparison to normal tissues were observed in the Human Protein Atlas Database [89,90] for: Breast carcinoma (BC); Prostate carcinoma (PC); Colorectal carcinoma (CC) and Lung carcinoma (LC). Protein expression levels are indicated by + (high/medium), - (low or not detected), \* (variable), ? (not analysed). Correlations between perturbations and expression data indicated in yellow.

Single Perturbations	Model adhesion phenotypes				Protein Expression of Carcinomas			
	F=0 C=2	F=1 C=0	F=1 C=2	F=1 C=1	BC	PC	LC	CC
<b>MET E1</b>	L	M	G	N	+	+	+	+
<b>EGFR E1</b>	L	M	G	N	-	-	-	*
<b>RAP1 KO</b>	L	M	N	N	?	?	?	?
<b>SRC E1</b>	L	M	N	N	-	+	-	+
<b>CDH1 KO</b>	L	M	N	N	*	*	+	+
<b>TGFb-TGFbR KO</b>	M	L	N	G	*	*	*	*
<b>SMAD KO</b>	M	L	N	G	*	*	*	*
<b>MEK KO</b>	M	L	N	N	+	+	-	-
<b>ERK KO</b>	M	L	N	N	+	+	+	+
<b>TCF_LEF KO</b>	M	L	N	G	+	+	+	+
<b>SNAI KO</b>	M	L	N	G	+	+	+	+
<b>ZEB KO</b>	M	L	N	G	-	-	-	-
<b>ILK KO</b>	M	L	N	G	-	-	-	-
<b>RPTP E1</b>	M	L	N	N	-	-	-	+
<b>FAT4 E1</b>	M	L	N	G	+	+	+	+
<b>LATS E1</b>	M	L	N	G	+	+	-	+
<b>YAP_TAZ KO</b>	M	L	N	G	-	-	-	-
<b>miR200 E1</b>	M	L	N	G	?	?	?	?
<b>CDH1 E1</b>	M	L	N	G	*	+	*	+
<b>MET KO</b>	M	M	N	G	+	+	+	+
<b>Dest_Complex E1</b>	M	M	N	G	?	?	?	?
<b>p120 E1</b>	M	M	N	G	*	*	*	*
<b>b-cat_nuc KO</b>	M	M	N	G	*	*	*	*

**Table S9. Comparison of the effects of double perturbations on model components with the expression of these proteins in human carcinomas.** The gains of a phenotype in comparison to the unperturbed model phenotypes is identified by the letter **G** and the loss of a phenotype by the letter **L**. Effects that favours or counteract tumour invasion are indicated in red or green, respectively. Perturbations that maintain phenotypes are indicated by the letter **M** and the ones that are not able to cause a gain are indicated by the letter **N**. Focal adhesion dynamics is abbreviated by **F** and cell-cell adhesion by **C**. Protein expression levels of carcinomas in comparison to normal issues were observed in the Human Protein Atlas Database [89,90] for: Breast carcinoma (BC); Prostate carcinoma (PC); Colorectal carcinoma (CC) and Lung carcinoma (LC). Protein expression levels are indicated by + (high/medium), - (low or not detected), \* (variable), ? (not analysed). Correlations between perturbations and expression data indicated in yellow.

Double Perturbations (Pi, Pj)	Model adhesion phenotypes				Protein (Pi) Expression				Protein (Pj) Expression			
	F=0 C=2	F=1 C=0	F=1 C=2	F=1 C=1	BC	PC	LC	CC	BC	PC	LC	CC
<b>NOTCH E1, RPTP KO</b>	L	M	N	N	+	+	+	+	-	-	-	+
<b>CSL E1, RPTP KO</b>	L	M	N	N	?	?	?	?	-	-	-	+
<b>TGFb-TGFbR E1, RPTP KO</b>	L	M	N	N	*	*	*	*	-	-	-	+
<b>JNK E1, RPTP KO</b>	L	M	N	N	*	*	*	*	-	-	-	+
<b>CK1 E1, RPTP KO</b>	L	M	N	N	-	-	-	+	-	-	-	+
<b>TCF_LEF E1, RPTP KO</b>	L	M	N	N	+	+	+	+	-	-	-	+
<b>CD44 E1, RPTP KO</b>	L	M	N	N	*	*	*	*	-	-	-	+
<b>ZEB E1, RPTP KO</b>	L	M	N	N	-	-	-	-	-	-	-	+
<b>SMAD E1, SNAI E1</b>	L	M	N	N	*	*	*	*	+	+	+	+
<b>PI3K E1, TCF_LEF E1</b>	L	M	N	N	+	+	+	+	+	+	+	+
<b>TCF_LEF E1, LATS KO</b>	L	M	N	N	+	+	+	+	+	+	-	+
<b>TCF_LEF E1, YAP_TAZ E1</b>	L	M	N	N	+	+	+	+	-	-	-	-
<b>SNAI E1, ZEB E1</b>	L	M	N	N	+	+	+	+	-	-	-	-
<b>MET KO, EGFR KO</b>	M	L	N	N	+	+	+	+	-	-	-	*
<b>MET KO, Wnt-Frizzled KO</b>	M	L	N	N	+	+	+	+	?	?	?	?
<b>MET KO, CK1 KO</b>	M	L	N	N	+	+	+	+	-	-	-	*
<b>MET KO, b-Cat_nuc KO</b>	M	L	N	N	+	+	+	+	*	*	*	*
<b>MET KO, Dest_Complex E1</b>	M	L	N	N	+	+	+	+	?	?	?	?
<b>MET KO, p120 E1</b>	M	L	N	N	+	+	+	+	*	*	*	*
<b>AKT KO, b-Cat_nuc KO</b>	M	L	N	N	+	+	*	+	*	*	*	*
<b>AKT KO, Dest_Complex E1</b>	M	L	N	N	+	+	*	+	?	?	?	?
<b>AKT KO, p120 E1</b>	M	L	N	N	+	+	*	+	*	*	*	*
<b>MET E1, p120 KO</b>	M	M	N	G	+	+	+	++	*	*	*	*
<b>EGFR E1, p120 KO</b>	M	M	N	G	-	-	-	*	*	*	*	*

**Table S10. Impact of perturbations on model interactions in terms of the reachability of adhesion phenotypes under chronic inflammation and RPTPL signals.** The probabilities of reaching adhesion phenotypes were estimated by simulation in GINsim using the method described in section 2.4.3. The letters C and F means cell-cell adhesion and Focal adhesion, respectively. The correspondent phenotypes are abbreviated by the letters E (Epithelial), H (Hybrid), A (amoeboid-like), and M (Mesenchymal). All simulations started from all state space and with the inputs that describe the chronic inflammation conditions (IL6=ECM=HGF=1 & ROS=EGF=0) with high RPTPL signals (RPTP=1 & DELTA=FAT4L=0).

Perturbations on model interactions	Reachable Adhesion Phenotypes
Unperturbed	F = 0 & C = 2 (E, p = 1)
Ectopic activation of SRC –  RPTP	F = 1 & C = 0 (M, p = 1)
KO of RPTP → p120	F = 0 & C = 2 (E, p = 0.9) F = 0 & C = 0 (A, p = 0.1)
KO of RPTP → b-Cat_mem	F = 0 & C = 1 (I, p = 1)
KO of RPTP –  EGFR	F = 1 & C = 2 (H, p = 1)
KO of RPTP –  MET	F = 1 & C = 2 (H, p = 0.8) F = 1 & C = 0 (M, p = 0.2)
KO of RPTP –  MET KO of RPTP –  EGFR	F = 1 & C = 2 (H, p = 0.9) F = 1 & C = 0 (M, p = 0.1)
KO of RPTP –  MET KO of RPTP → b-Cat_mem	F = 1 & C = 1 (H, p = 0.8) F = 1 & C = 0 (M, p = 0.2)
KO of RPTP → p120 KO of RPTP → b-Cat_mem	F = 0 & C = 0 (A, p = 1)
KO of RPTP –  MET KO of RPTP → b-Cat_mem KO of RPTP → p120	F = 1 & C = 0 (M, p = 1)
KO of RPTP –  EGFR KO of RPTP → b-Cat_mem KO of RPTP → p120	F = 0 & C = 0 (A, p = 1)
Ectopic activation of MET → FAK_SRC	F = 1 & C = 2 (H, p = 0.8) F = 0 & C = 2 (E, p = 0.2)
Ectopic activation of EGFR → FAK_SRC	F = 1 & C = 2 (H, p = 0.8) F = 0 & C = 2 (E, p = 0.2)
Ectopic activation of MET → FAK_SRC Ectopic activation of EGFR → FAK_SRC	F = 1 & C = 2 (H, p = 0.8) F = 0 & C = 2 (E, p = 0.2)

**Table S11. Impact of perturbations on model interactions (Hippo signalling) in terms of the reachability of adhesion phenotypes under chronic inflammation and ectopic FAT4 signals.** The probabilities of reaching adhesion phenotypes were estimated by simulation in GINsim using the method described in section 2.4.3. The letters C and F means cell-cell adhesion and Focal adhesion, respectively. The correspondent phenotypes are abbreviated by the letters E (Epithelial), H (Hybrid), A (amoeboid-like), and M (Mesenchymal). All simulations started from all state space and with the inputs that describe the chronic inflammation conditions (IL6=ECM=HGF=1 & ROS=EGF=0) with high RPTPL signals (RPTP=1 & DELTA=FAT4L=0).

Perturbations on model interactions	Reachable Adhesion Phenotypes
FAT4 ectopic activation (FAT4 E1)	F = 1 & C = 1 (H, p = 1)
FAT4 E1 KO of FAT4 → LATS	F = 1 & C = 0 (M, p = 1)
FAT4 E1 KO of LATS → YAP_TAZ	F = 1 & C = 0 (M, p = 1)
FAT4 E1 Ectopic activation of YAP_TAZ → b-Cat_nuc	F = 1 & C = 1 (H, p = 1)
FAT4 E1 Ectopic activation of YAP_TAZ → SMAD	F = 1 & C = 1 (H, p = 0.9) F = 1 & C = 0 (M, p = 0.1)
FAT4 E1 Ectopic activation of YAP_TAZ → DVL	F = 1 & C = 1 (H, p = 1)
FAT4 E1 Ectopic activation of YAP_TAZ → SMAD Ectopic activation of YAP_TAZ → DVL	F = 1 & C = 1 (H, p = 0.9) F = 1 & C = 0 (M, p = 0.1)
FAT4 E1 Ectopic activation of YAP_TAZ → SMAD Ectopic activation of YAP_TAZ → b-Cat_nuc	F = 1 & C = 0 (M, p = 1)

## 6. References

1. Savagner P. Epithelial-mesenchymal transitions: from cell plasticity to concept elasticity. Curr Top Dev Biol. 2015;112: 273–300. doi:10.1016/bs.ctdb.2014.11.021
2. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression. J Mammary Gland Biol Neoplasia. 2010;15: 117–34. doi:10.1007/s10911-010-9178-9
3. Kalluri R, Weinberg R a. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119: 1420–1428. doi:10.1172/JCI39104.1420
4. Tsai JH, Yang J. Epithelial-mesenchymal plasticity in carcinoma metastasis. Genes Dev. 2013;27: 2192–206. doi:10.1101/gad.225334.113



5. Steinestel K, Eder S, Schrader AJ, Steinestel J. Clinical significance of epithelial-mesenchymal transition. *Clin Transl Med.* 2014;3: 1–13.
6. Tsuji T, Ibaragi S, Hu G. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res.* 2009;69: 7135–9. doi:10.1158/0008-5472.CAN-09-1618
7. Klymkowsky MW, Savagner P. Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. *Am J Pathol.* 2009;174: 1588–93. doi:10.2353/ajpath.2009.080545
8. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Onuchic JN, et al. Implications of the hybrid epithelial/mesenchymal phenotype in metastasis. 2015;5: 1–19. doi:10.3389/fonc.2015.00155
9. Garg M. Epithelial, mesenchymal and hybrid epithelial/mesenchymal phenotypes and their clinical relevance in cancer metastasis. *Expert Rev Mol Med.* 2017;19: e3. doi:10.1017/erm.2017.6
10. Kawauchi T. Cell adhesion and its endocytic regulation in cell migration during neural development and cancer metastasis. *Int J Mol Sci. Multidisciplinary Digital Publishing Institute (MDPI);* 2012;13: 4564–90. doi:10.3390/ijms13044564
11. Clark AG, Vignjevic DM. Modes of cancer cell invasion and the role of the microenvironment. *Curr Opin Cell Biol.* 2015;36: 13–22. doi:10.1016/j.ceb.2015.06.004
12. Canel M, Serrels A, Frame MC, Brunton VG. E-cadherin–integrin crosstalk in cancer invasion and metastasis. *J Cell Sci.* 2013;126: 393–401. doi:10.1242/jcs.100115
13. Nagano M, Hoshino D, Koshikawa N, Akizawa T, Seiki M. Turnover of focal adhesions and cancer cell migration. *Int J Cell Biol.* 2011;2012: 1–12. doi:10.1155/2012/310616
14. Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Biol. Nature Publishing Group;* 2009;10: 445–457. doi:10.1038/nrm2720
15. Friedl P, Alexander S. Cancer invasion and the microenvironment: Plasticity and reciprocity. *Cell. Elsevier Inc.;* 2011;147: 992–1009. doi:10.1016/j.cell.2011.11.016
16. Li H, Xu F, Li S, Zhong A, Meng X, Lai M. The tumor microenvironment: An irreplaceable element of tumor budding and epithelial-mesenchymal transition-mediated cancer metastasis. *Cell Adh Migr. Taylor & Francis;* 2016;10: 434–446. doi:10.1080/19336918.2015.1129481
17. Qin J-H, Wang L, Li Q-L, Liang Y, Ke Z-Y, Wang R-A. Epithelial-mesenchymal transition as strategic microenvironment mimicry for cancer cell survival and immune escape? [Internet]. *Genes & Diseases.* 2017. pp. 16–18. doi:10.1016/j.gendis.2016.10.001
18. Krakhmal N V, Zavyalova M V, Denisov E V, Vtorushin S V, Perelmuter VM. Cancer invasion: Patterns and mechanisms [Internet]. *Acta Naturae. Park Media;* 2015. pp. 17–28. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26085941>

19. Gao D, Vahdat LT, Wong S, Chang JC, Mittal V. Microenvironmental regulation of epithelial-mesenchymal transitions in cancer. *Cancer Res.* NIH Public Access; 2012;72: 4883–9. doi:10.1158/0008-5472.CAN-12-1223
20. Huang B, Jolly MK, Lu M, Tsarfaty I, Ben-Jacob E, Onuchic JN. Modeling the Transitions between Collective and Solitary Migration Phenotypes in Cancer Metastasis. *Sci Rep.* 2015;5: 17379. doi:10.1038/srep17379
21. Place AE, Jin Huh S, Polyak K. The microenvironment in breast cancer progression: biology and implications for treatment. *Breast Cancer Res. BioMed Central*; 2011;13: 227. doi:10.1186/bcr2912
22. Landskron G, De la Fuente M, Thuwajit P, Thuwajit C, Hermoso MA. Chronic inflammation and cytokines in the tumor microenvironment. *J Immunol Res.* 2014;2014: 149185. doi:10.1155/2014/149185
23. Gilkes DM, Semenza GL, Wirtz D. Hypoxia and the extracellular matrix: drivers of tumour metastasis. *Nat Rev Cancer.* 2014;14: 430–439. doi:10.1038/nrc3726
24. Deneff C. Contact-dependent Signaling. *Cell Commun Insights.* 2014; 1–11. doi:10.4137/CCI.s12484.TYPE
25. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* Nature Publishing Group; 2014;15: 178–96. doi:10.1038/nrm3758
26. Nieto MA, Cano A. The epithelial-mesenchymal transition under control: Global programs to regulate epithelial plasticity. *Semin Cancer Biol.* Elsevier Ltd; 2012;22: 361–368. doi:10.1016/j.semcancer.2012.05.003
27. De Craene B, Berx G. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer.* Nature Publishing Group; 2013;13: 97–110. doi:10.1038/nrc3447
28. Kandath C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013;502: 333–9. doi:10.1038/nature12634
29. Somvanshi PR, Venkatesh K V. A conceptual review on systems biology in health and diseases: from biological networks to modern therapeutics. *Syst Synth Biol.* Springer; 2014;8: 99–116. doi:10.1007/s11693-013-9125-3
30. Rateitschak K, Kaderali L, Wolkenhauer O, Jaster R. Autocrine TGF- $\beta$ /ZEB/microRNA-200 signal transduction drives epithelial-mesenchymal transition: Kinetic models predict minimal drug dose to inhibit metastasis. *Cell Signal.* 2016;28: 861–870. doi:10.1016/j.cellsig.2016.03.002
31. Steinway SN, Zanudo JGT, Ding W, Rountree CB, Feith DJ, Loughran TP, et al. Network modeling of TGF $\beta$  signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint Sonic hedgehog and Wnt pathway activation. *Cancer Res.* 2014;74: 5963–5977. doi:10.1158/0008-5472.CAN-14-0225

32. Cohen DPA, Martignetti L, Robine S, Barillot E, Zinovyev A, Calzone L. Mathematical Modelling of Molecular Pathways Enabling Tumour Cell Invasion and Migration. *PLoS Comput Biol*. 2015;11. doi:10.1371/journal.pcbi.1004571
33. Howard S, Deroo T, Fujita Y, Itasaki N. A positive role of cadherin in Wnt/ $\beta$ -catenin signalling during epithelial-mesenchymal transition. *PLoS One*. 2011;6: e23899. doi:10.1371/journal.pone.0023899
34. Zhang J, Tian X-JX-J, Zhang H, Teng Y, Li R, Bai F, et al. TGF $\beta$  -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci Signal*. 2014;7: ra91-ra91. doi:10.1126/scisignal.2005304
35. Steinway SN, Zañudo JGT, Michel PJ, Feith DJ, Loughran TP, Albert R. Combinatorial interventions inhibit TGF $\beta$ -driven epithelial-to-mesenchymal transition and support hybrid cellular phenotypes. *npj Syst Biol Appl*. 2015;1: 1–12. doi:10.1038/npsba.2015.14
36. Le Novère N. Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet*. Europe PMC Funders; 2015;16: 146–58. doi:10.1038/nrg3885
37. Grieco L, Calzone L, Bernard-Pierrot I, Radvanyi F, Kahn-Perlès B, Thieffry D. Integrative modelling of the influence of MAPK network on cancer cell fate decision. *PLoS Comput Biol*. 2013;9: e1003286. doi:10.1371/journal.pcbi.1003286
38. Flobak Å, Baudot A, Remy E, Thommesen L, Thieffry D, Kuiper M, et al. Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling. *PLoS Comput Biol*. 2015;11. doi:10.1371/journal.pcbi.1004426
39. Remy E, Rebouissou S, Chaouiya C, Zinovyev A, Radvanyi F, Calzone L. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. *Cancer Res*. 2015;75: 4042–4052. doi:10.1158/0008-5472.CAN-15-0602
40. Xu Y, Fisher GJ. Receptor type protein tyrosine phosphatases (RPTPs) - roles in signal transduction and human disease. *J Cell Commun Signal*. Springer; 2012;6: 125–38. doi:10.1007/s12079-012-0171-5
41. Katoh M. Function and cancer genomics of FAT family genes (Review). *Int J Oncol*. 2012;41: 1913–8. doi:10.3892/ijo.2012.1669
42. Hertog J den, Östman A, Böhmer F-D. Protein tyrosine phosphatases: regulatory mechanisms. *FEBS J*. Blackwell Publishing Ltd; 2008;275: 831–847. doi:10.1111/j.1742-4658.2008.06247.x
43. Mao Y, Francis-West P, Irvine KD. Fat4/Dchs1 signaling between stromal and cap mesenchyme cells influences nephrogenesis and ureteric bud branching. *Development*. 2015;142: 2574–85. doi:10.1242/dev.122630
44. Mohebiany AN, Nikolaienko RM, Bouyain S, Harroch S. Receptor-type tyrosine phosphatase ligands: looking for the needle in the haystack. *FEBS J*. 2013;280: 388–400. doi:10.1111/j.1742-4658.2012.08653.x

45. Sap J, Jiang YP, Friedlander D, Grumet M, Schlessinger J. Receptor tyrosine phosphatase R-PTP-kappa mediates homophilic binding. *Mol Cell Biol. American Society for Microbiology (ASM)*; 1994;14: 1–9. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8264577>
46. Ostman A, Yang Q, Tonks NK. Expression of DEP-1, a receptor-like protein-tyrosine-phosphatase, is enhanced with increasing cell density. *Proc Natl Acad Sci U S A. National Academy of Sciences*; 1994;91: 9680–4. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7937872>
47. C. Chaouiya, A. Naldi DT. Logical Modelling of Gene Regulatory Networks with GINsim. *Methods in Molecular Biology*,. 1st ed. Springer; 2012. pp. 463–479. Available: [http://dx.doi.org/10.1007/978-1-61779-361-5\\_23](http://dx.doi.org/10.1007/978-1-61779-361-5_23)
48. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res. Oxford University Press*; 2017;45: D353–D361. doi:10.1093/nar/gkw1092
49. Kuperstein I, Bonnet E, Nguyen H-AH, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. *Oncogenesis*. 2015;4: 1–14. doi:10.1038/oncsis.2015.19
50. Kuperstein I, Cohen DPA, Pook S, Viara E, Calzone L, Barillot E, et al. NaviCell: a web-based environment for navigation, curation and maintenance of large molecular interaction maps. *BMC Syst Biol*. 2013;7: 100. doi:10.1186/1752-0509-7-100
51. Li G-W, Xie XS. Central dogma at the single-molecule level in living cells. *Nature*. 2011;475: 308–315. doi:10.1038/nature10315
52. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473: 337–342. doi:10.1038/nature10098
53. Luo D, Smith SW, Anderson BD. Kinetics and Mechanism of the Reaction of Cysteine and Hydrogen Peroxide in Aqueous Solution. *J Pharm Sci*. 2005;94: 304–316. doi:10.1002/jps.20253
54. Antunes F, Brito PM. Quantitative biology of hydrogen peroxide signaling. *Redox Biol*. 2017; doi:10.1016/j.redox.2017.04.039
55. Blazek M, Santisteban TS, Zengerle R, Meier M. Analysis of fast protein phosphorylation kinetics in single cells on a microfluidic chip. *Lab Chip*. 2015;15: 726–34. doi:10.1039/c4lc00797b
56. Kleiman LB, Maiwald T, Conzelmann H, Lauffenburger DA, Sorger PK. Rapid phospho-turnover by receptor tyrosine kinases impacts downstream signaling and drug binding. *Mol Cell. NIH Public Access*; 2011;43: 723–37. doi:10.1016/j.molcel.2011.07.014
57. Northrup SH, Erickson HP. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc Natl Acad Sci U S A. National Academy of Sciences*; 1992;89: 3338–42. Available: <http://www.ncbi.nlm.nih.gov/pubmed/1565624>

58. Arhel N, Genovesio A, Kim K-A, Miko S, Perret E, Olivo-Marin J-C, et al. Quantitative four-dimensional tracking of cytoplasmic and nuclear HIV-1 complexes. *Nat Methods*. 2006;3: 817–824. doi:10.1038/nmeth928
59. Thomas R. Regulatory networks seen as asynchronous automata: A logical description. *J Theor Biol. Academic Press*; 1991;153: 1–23. doi:10.1016/S0022-5193(05)80350-9
60. Faure A, Naldi A, Chaouiya C, Thieffry D. Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics. Oxford University Press*; 2006;22: e124–e131. doi:10.1093/bioinformatics/btl210
61. Naldi A, Thieffry D, Chaouiya C. Decision Diagrams for the Representation and Analysis of Logical Models of Genetic Networks. *Computational Methods in Systems Biology*. 2007. pp. 233–244. Available: <http://aurelien.naldi.info/references/Naldi2007.pdf>
62. Béranguier D, Chaouiya C, Monteiro PT, Naldi a, Remy E, Thieffry D, et al. Dynamical modeling and analysis of large cellular regulatory networks. *Chaos*. 2013;23: 25114. doi:10.1063/1.4809783
63. Blau HM, Blakely BT. Plasticity of cell fate: Insights from heterokaryons. *Semin Cell Dev Biol*. 1999;10: 267–272. doi:10.1006/scdb.1999.0311
64. Zeisberg M, Neilson EG. Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest. American Society for Clinical Investigation*; 2009;119: 1429–37. doi:10.1172/JCI36183
65. Ding S, Zhang W, Xu Z, Xing C, Xie H, Guo H, et al. Induction of an EMT-like transformation and MET in vitro. *J Transl Med. BioMed Central*; 2013;11: 164. doi:10.1186/1479-5876-11-164
66. Zhu Y-T, Chen H-C, Chen S-Y, Tseng SCG. Nuclear p120 catenin unlocks mitotic block of contact-inhibited human corneal endothelial monolayers without disrupting adherent junctions. *J Cell Sci*. 2012;125: 3636–3648. doi:10.1242/jcs.103267
67. Davis M a., Ireton RC, Reynolds AB. A core function for p120-catenin in cadherin turnover. *J Cell Biol*. 2003;163: 525–534. doi:10.1083/jcb.200307111
68. Lickert H. Casein Kinase II Phosphorylation of E-cadherin Increases E-cadherin/beta -Catenin Interaction and Strengthens Cell-Cell Adhesion. *J Biol Chem*. 2000;275: 5090–5095. doi:10.1074/jbc.275.7.5090
69. Dupre-Crochet S, Figueroa A, Hogan C, Ferber EC, Bialucha CU, Adams J, et al. Casein kinase 1 is a novel negative regulator of E-cadherin-based cell-cell contacts. *Mol Cell Biol*. 2007;27: 3804–16. doi:10.1128/MCB.01590-06
70. Fujisaki T, Tanaka Y, Fujii K, Mine S, Saito K, Yamada S, et al. CD44 Stimulation Induces Integrin-mediated Adhesion of Colon Cancer Cell Lines to Endothelial Cells by Up-Regulation of Integrins and c-Met and Activation of Integrins. *Cancer Res*. 1999;59: 4427–4434. Available: <http://cancerres.aacrjournals.org/content/59/17/4427.long>
71. Behrens J, Vakaet L, Friis R, Winterhager E, Van Roy F, Mareel MM, et al. Loss of epithelial differentiation and gain of invasiveness correlates with tyrosine phosphorylation of the E-

- cadherin/beta-catenin complex in cells transformed with a temperature-sensitive v-SRC gene. *J Cell Biol.* 1993;120: 757–66. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8425900>
72. Avizienyte E, Wyke AW, Jones RJ, McLean GW, Westhoff MA, Brunton VG, et al. Src-induced de-regulation of E-cadherin in colon cancer cells requires integrin signalling. *Nat Cell Biol.* 2002;4: 632–8. doi:10.1038/ncb829
  73. Park HB, Golubovskaya V, Xu L, Yang X, Lee JW, Scully S, et al. Activated Src increases adhesion, survival and alpha2-integrin expression in human breast cancer cells. *Biochem J.* 2004;378: 559–67. doi:10.1042/BJ20031392
  74. Asgeirsson KS, Olafsdóttir K, Jónasson JG, Ogmundsdóttir HM. The effects of IL-6 on cell adhesion and e-cadherin expression in breast cancer. *Cytokine.* 1998;10: 720–8. doi:10.1006/cyto.1998.0349
  75. Cheng J-C, Chang H-M, Leung PCK. Egr-1 mediates epidermal growth factor-induced downregulation of E-cadherin expression via Slug in human ovarian cancer cells. *Oncogene.* Nature Publishing Group; 2012;32: 1041–1049. doi:10.1038/onc.2012.127
  76. Liang Q, Mohan RR, Chen L, Wilson SE. Signaling by HGF and KGF in corneal epithelial cells: Ras/MAP kinase and Jak-STAT pathways. *Invest Ophthalmol Vis Sci.* 1998;39: 1329–38. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9660480>
  77. Fujii K, Furukawa F, Matsuyoshi N. Ligand activation of overexpressed epidermal growth factor receptor results in colony dissociation and disturbed E-cadherin function in HSC-1 human cutaneous squamous carcinoma cells. *Exp Cell Res.* 1996;223: 50–62. doi:10.1006/excr.1996.0057
  78. Lu Z, Jiang G, Blume-jensen P, Hunter T. Epidermal Growth Factor-Induced Tumor Cell Invasion and Metastasis Initiated by Dephosphorylation and Downregulation of Focal Adhesion Kinase. *Society.* 2001;21: 4016–4031. doi:10.1128/MCB.21.12.4016
  79. Cheng JC, Leung PCK. Type I collagen down-regulates E-cadherin expression by increasing PI3KCA in cancer cells. *Cancer Lett.* Elsevier Ireland Ltd; 2011;304: 107–116. doi:10.1016/j.canlet.2011.02.008
  80. Grotegut S, von Schweinitz D, Christofori G, Lehenbre FF. Hepatocyte growth factor induces cell scattering through MAPK/Egr-1-mediated upregulation of Snail. *EMBO J.* 2006;25: 3534–3545. doi:10.1038/sj.emboj.7601213
  81. Farrell J, Kelly C, Rauch J, Kida K, García-Muñoz A, Monsefi N, et al. HGF Induces Epithelial-to-Mesenchymal Transition by Modulating the Mammalian Hippo/MST2 and ISG15 Pathways. *J Proteome Res.* 2014;13: 2874–2886. doi:10.1021/pr5000285
  82. Liu Z-X, Yu CF, Nickel C, Thomas S, Cantley LG. Hepatocyte Growth Factor Induces ERK-dependent Paxillin Phosphorylation and Regulates Paxillin-Focal Adhesion Kinase Association. *J Biol Chem.* 2002;277: 10452–10458. doi:10.1074/jbc.M107551200

83. Del Valle-Pérez B, Casagolda D, Lugilde E, Valls G, Codina M, Dave N, et al. Wnt controls the transcriptional activity of Kaiso through CK1 $\epsilon$ -dependent phosphorylation of p120-catenin. *J Cell Sci.* 2011;124: 2298–309. doi:10.1242/jcs.082693
84. Pearce JD, Edwards MS, Craven TE, English WP, Mondt MM, Reavis SW, et al. Renal duplex parameters, blood pressure, and renal function in elderly people. *Am J Kidney Dis.* 2005;45: 842–50. Available: <http://www.ncbi.nlm.nih.gov/pubmed/15861349>
85. Hiemer SE, Szymaniak AD, Varelas X. The transcriptional regulators TAZ and YAP direct transforming growth factor  $\beta$ -induced tumorigenic phenotypes in breast cancer cells. *J Biol Chem.* 2014;289: 13461–74. doi:10.1074/jbc.M113.529115
86. Edlund S, Landström M, Heldin C-H, Aspenström P. Transforming growth factor-beta-induced mobilization of actin cytoskeleton requires signaling by small GTPases Cdc42 and RhoA. *Mol Biol Cell.* 2002;13: 902–14. doi:10.1091/mbc.01-08-0398
87. Kaimori A, Potter J, Kaimori J-Y, Wang C, Mezey E, Koteish A. Transforming growth factor-beta1 induces an epithelial-to-mesenchymal transition state in mouse hepatocytes in vitro. *J Biol Chem.* 2007;282: 22089–101. doi:10.1074/jbc.M700998200
88. Chan H-L, Chou H-C, Duran M, Gruenewald J, Waterfield MD, Ridley A, et al. Major role of epidermal growth factor receptor and Src kinases in promoting oxidative stress-dependent loss of adhesion and apoptosis in epithelial cells. *J Biol Chem.* 2010;285: 4307–18. doi:10.1074/jbc.M109.047027
89. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science (80- ).* 2015;347: 1260419–1260419. doi:10.1126/science.1260419
90. Uhlén M, Björling E, Agaton C, Szgyarto CA-K, Amini B, Andersen E, et al. A Human Protein Atlas for Normal and Cancer Tissues Based on Antibody Proteomics. *Mol Cell Proteomics.* 2005;4: 1920–1932. doi:10.1074/mcp.M500279-MCP200
91. Espinoza I, Miele L. Deadly crosstalk: Notch signaling at the intersection of EMT and cancer stem cells. *Cancer Lett.* 2013;341: 41–45. doi:10.1016/j.canlet.2013.08.027
92. Legate KR, Wickström S a, Fässler R, Fa R, Wickstro S a. Genetic and cell biological analysis of integrin outside-in signaling. *Genes Dev.* 2009;23: 397–418. doi:10.1101/gad.1758709
93. Kim SH, Turnbull J, Guimond S. Extracellular matrix and cell signalling: The dynamic cooperation of integrin, proteoglycan and growth factor receptor. *J Endocrinol.* 2011;209: 139–151. doi:10.1530/JOE-10-0377
94. Eberwein P, Laird D, Schulz S, Reinhard T, Steinberg T, Tomakidi P. Modulation of focal adhesion constituents and their down-stream events by EGF: On the cross-talk of integrins and growth factor receptors. *Biochim Biophys Acta - Mol Cell Res.* 2015;1853: 2183–2198. doi:10.1016/j.bbamcr.2015.06.004

95. Kim M, Jho E-H. Cross-talk between Wnt/ $\beta$ -catenin and Hippo signaling pathways: a brief review. *BMB Rep.* 2014;47: 540–5. Available: <http://www.ncbi.nlm.nih.gov/pubmed/25154721>
96. Varelas X, Miller BW, Sopko R, Song S, Gregorieff A, Fellouse FA, et al. The Hippo pathway regulates Wnt/beta-catenin signaling. *Dev Cell.* 2010;18: 579–91. doi:10.1016/j.devcel.2010.03.007
97. Vicente-Manzanares M, Horwitz AR. Adhesion dynamics at a glance. *J Cell Sci.* Company of Biologists; 2011;124: 3923–3927. doi:10.1242/jcs.095653
98. Adams CL, Chen YT, Smith SJ, Nelson WJ. Mechanisms of epithelial cell-cell adhesion and cell compaction revealed by high-resolution tracking of E-cadherin-green fluorescent protein. *J Cell Biol.* 1998;142: 1105–19. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9722621>
99. Thoreson MA, Anastasiadis PZ, Daniel JM, Ireton RC, Wheelock MJ, Johnson KR, et al. Selective uncoupling of p120(ctn) from E-cadherin disrupts strong adhesion. *J Cell Biol.* 2000;148: 189–202. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10629228>
100. Pacquelet A, Lin L, Rorth P. Binding site for p120/beta-catenin is not required for Drosophila E-cadherin function in vivo. *J Cell Biol.* 2003;160: 313–319. doi:10.1083/jcb.200207160
101. Le Bras GF, Taubenslag KJ, Andl CD. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adh Migr.* Taylor & Francis; 2012;6: 365–73. doi:10.4161/cam.21326
102. Leopold PL, Vincent J, Wang H. A comparison of epithelial-to-mesenchymal transition and re- epithelialization. *Semin Cancer Biol.* 2012;22: 471–483. doi:10.1016/j.semcancer.2012.07.003
103. Larue L, Bellacosa A. Epithelial-mesenchymal transition in development and cancer: role of phosphatidylinositol 3' kinase/AKT pathways. *Oncogene.* 2005;24: 7443–54. doi:10.1038/sj.onc.1209091
104. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. *Oncotarget.* Impact Journals, LLC; 2016;7: 27067–27084. doi:10.18632/oncotarget.8166
105. Huang B, Lu M, Jolly MK, Tsarfaty I, Onuchic J, Ben-Jacob E. The three-way switch operation of Rac1/RhoA GTPase-based circuit controlling amoeboid-hybrid-mesenchymal transition. *Sci Rep.* 2014;4: 6449. doi:10.1038/srep06449
106. Panková K, Rösel D, Novotný M, Brábek J. The molecular mechanisms of transition between mesenchymal and amoeboid invasiveness in tumor cells. *Cell Mol Life Sci.* Springer; 2010;67: 63–71. doi:10.1007/s00018-009-0132-1
107. Silver DL, Geisbrecht ER, Montell DJ. Requirement for JAK/STAT signaling throughout border cell migration in Drosophila. *Development.* 2005;132: 3483–92. doi:10.1242/dev.01910



108. Xu Y, Shao Y, Voorhees JJ, Fisher GJ. Oxidative Inhibition of Receptor-type Protein-tyrosine Phosphatase  $\kappa$  by Ultraviolet Irradiation Activates Epidermal Growth Factor Receptor in Human Keratinocytes. *J Biol Chem*. 2006;281: 27389–27397. doi:10.1074/jbc.M602355200
109. Godfrey R, Arora D, Bauer R, Stopp S, Müller JP, Heinrich T, et al. Cell transformation by FLT3 ITD in acute myeloid leukemia involves oxidative inactivation of the tumor suppressor protein-tyrosine phosphatase DEP-1/ PTPRJ. *Blood*. 2012;119: 4499–511. doi:10.1182/blood-2011-02-336446
110. Shao S, Zhao X, Zhang X, Luo M, Zuo X, Huang S, et al. Notch1 signaling regulates the epithelial?mesenchymal transition and invasion of breast cancer in a Slug-dependent manner. *Mol Cancer*. 2015;14: 28. doi:10.1186/s12943-015-0295-3
111. Irby RB, Yeatman TJ. Role of Src expression and activation in human cancer. *Oncogene*. 2000;19: 5636–5642. Available: <http://www.nature.com/onc/journal/v19/n49/pdf/1203912a.pdf>
112. Neuzillet C, Tijeras-Raballand A, Cohen R, Cros J, Faivre S, Raymond E, et al. Targeting the TGF $\beta$  pathway for cancer therapy. *Pharmacol Ther*. 2014;142: 316–338. doi:10.1016/j.pharmthera.2014.11.001
113. Mao Y, Kuta A, Crespo-Enriquez I, Whiting D, Martin T, Mulvaney J, et al. Dchs1-Fat4 regulation of polarized cell behaviours during skeletal morphogenesis. *Nat Commun*. 2016;7: 11469. doi:10.1038/ncomms11469
114. Vadlamudi RK, Sahin AA, Adam L, Wang RA, Kumar R. Heregulin and HER2 signaling selectively activates c-Src phosphorylation at tyrosine 215. *FEBS Lett*. 2003;543: 76–80. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12753909>
115. Schaller MD, Hildebrand JD, Parsons JT. Complex formation with focal adhesion kinase: A mechanism to regulate activity and subcellular localization of Src kinases. *Mol Biol Cell*. American Society for Cell Biology; 1999;10: 3489–505. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10512882>
116. Mitra SK, Schlaepfer DD. Integrin-regulated FAK–Src signaling in normal and cancer cells. *Curr Opin Cell Biol*. 2006;18: 516–523. doi:10.1016/j.ceb.2006.08.011
117. Nelson WJ. Regulation of cell-cell adhesion by the cadherin-catenin complex. *Biochem Soc Trans*. NIH Public Access; 2008;36: 149–55. doi:10.1042/BST0360149
118. Yamada S, Pokutta S, Drees F, Weis WI, Nelson WJ. Deconstructing the Cadherin-Catenin-Actin Complex. *Cell*. 2005;123: 889–901. doi:10.1016/j.cell.2005.09.020
119. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011;144: 646–674. doi:10.1016/j.cell.2011.02.013
120. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett*. No longer published by Elsevier; 2009;583: 3966–3973. doi:10.1016/J.FEBSLET.2009.10.036

121. Prelich G. Gene overexpression: uses, mechanisms, and interpretation. *Genetics*. Genetics Society of America; 2012;190: 841–54. doi:10.1534/genetics.111.136911
122. Du Y, Grandis JR. Receptor-type protein tyrosine phosphatases in cancer. *Chin J Cancer*. BioMed Central; 2015;34: 61–9. doi:10.5732/cjc.014.10146
123. Holsinger LJ, Ward K, Duffield B, Zachwieja J, Jallal B. The transmembrane receptor protein tyrosine phosphatase DEP1 interacts with p120(ctn). *Oncogene*. Nature Publishing Group; 2002;21: 7067–76. doi:10.1038/sj.onc.1205858
124. Palka HL, Park M, Tonks NK. Hepatocyte Growth Factor Receptor Tyrosine Kinase Met Is a Substrate of the Receptor Protein-tyrosine Phosphatase DEP-1. *J Biol Chem*. 2003;278: 5728–5735. doi:10.1074/jbc.M210656200
125. Xu Y, Tan L-J, Grachtchouk V, Voorhees JJ, Fisher GJ. Receptor-type Protein-tyrosine Phosphatase- $\kappa$  Regulates Epidermal Growth Factor Receptor Function. *J Biol Chem*. 2005;280: 42694–42700. doi:10.1074/jbc.M507722200
126. Hirsch HA, Iliopoulos D, Joshi A, Zhang Y, Jaeger SA, Bulyk M, et al. A transcriptional signature and common gene networks link cancer with lipid metabolism and diverse human diseases. *Cancer Cell*. 2010;17: 348–61. doi:10.1016/j.ccr.2010.01.022
127. Garcia R, Bowman TL, Niu G, Yu H, Minton S, Muro-Cacho CA, et al. Constitutive activation of Stat3 by the Src and JAK tyrosine kinases participates in growth regulation of human breast carcinoma cells. *Oncogene*. 2001;20. doi:10.1038/sj.onc.1204349
128. Xu Y, Baker D, Quan T, Baldassare JJ, Voorhees JJ, Fisher GJ. Receptor Type Protein Tyrosine Phosphatase-Kappa Mediates Cross-Talk between Transforming Growth Factor-Beta and Epidermal Growth Factor Receptor Signaling Pathways in Human Keratinocytes. *Mol Biol Cell*. American Society for Cell Biology; 2010;21: 29–35. doi:10.1091/mbc.E09-08-0710
129. Wang SE, Wu FY, Shin I, Qu S, Arteaga CL. Transforming growth factor {beta} (TGF-{beta})-Smad target gene protein tyrosine phosphatase receptor type kappa is required for TGF-{beta} function. *Mol Cell Biol*. American Society for Microbiology (ASM); 2005;25: 4703–15. doi:10.1128/MCB.25.11.4703-4715.2005
130. Ito T, Taniguchi H, Fukagai K, Okamuro S, Kobayashi A. Inhibitory mechanism of FAT4 gene expression in response to actin dynamics during Src-induced carcinogenesis. Rishi A, editor. *PLoS One*. 2015;10: e0118336. doi:10.1371/journal.pone.0118336
131. Gill JG, Langer EM, Lindsley RC, Cai M, Murphy TL, Kyba M, et al. Snail and the microRNA-200 Family Act in Opposition to Regulate Epithelial-to-Mesenchymal Transition and Germ Layer Fate Restriction in Differentiating ESCs. *Stem Cells*. 2011;29: 764–776. doi:10.1002/stem.628
132. Bracken CP, Gregory PA, Kolesnikoff N, Bert AG, Wang J, Shannon MF, et al. A Double-Negative Feedback Loop between ZEB1-SIP1 and the microRNA-200 Family Regulates Epithelial-Mesenchymal Transition. *Cancer Res*. 2008;68: 7846–7854. doi:10.1158/0008-5472.CAN-08-1942

133. Eastman Q, Grosschedl R. Regulation of LEF-1/TCF transcription factors by Wnt and other signals. *Curr Opin Cell Biol.* 1999;11: 233–240. doi:10.1016/S0955-0674(99)80031-3
134. MacDonald BT, Tamai K, He X. Wnt/ $\beta$ -Catenin Signaling: Components, Mechanisms, and Diseases. *Dev Cell.* 2009;17: 9–26. doi:10.1016/j.devcel.2009.06.016
135. Ishii N, Harada N, Joseph EW, Ohara K, Miura T, Sakamoto H, et al. Enhanced inhibition of ERK signaling by a novel allosteric MEK inhibitor, CH5126766, that suppresses feedback reactivation of RAF activity. *Cancer Res.* 2013;73: 4050–4060. doi:10.1158/0008-5472.CAN-12-3937
136. Tran KA, Cheng MY, Mitra A, Ogawa H, Shi VY, Olney LP, et al. MEK inhibitors and their potential in the treatment of advanced melanoma: the advantages of combination therapy. *Drug Des Devel Ther.* Dove Press; 2016;10: 43–52. doi:10.2147/DDDT.S93545
137. Maroun CR, Rowlands T. The Met receptor tyrosine kinase: A key player in oncogenesis and drug resistance. 2014; doi:10.1016/j.pharmthera.2013.12.014
138. Mayer IA, Arteaga CL. The PI3K/AKT Pathway as a Target for Cancer Treatment. *Annu Rev Med.* 2015;67: 11–28. doi:10.1146/annurev-med-062913-051343
139. Hartsock A, Nelson WJ. Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta - Biomembr.* NIH Public Access; 2008;1778: 660–669. doi:10.1016/j.bbamem.2007.07.012
140. Bolós V, Gasent JM, López-Tarruella S, Grande E. The dual kinase complex FAK-Src as a promising therapeutic target in cancer. *Onco Targets Ther.* 2010;3: 83–97. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20616959>
141. Retta SF, Balzac F, Avolio M. Rap1: A turnabout for the crosstalk between cadherins and integrins. *Eur J Cell Biol.* 2006;85: 283–293. doi:10.1016/j.ejcb.2005.09.007
142. Xiao K, Oas RG, Chiasson CM, Kowalczyk AP. Role of p120-catenin in cadherin trafficking. *Biochim Biophys Acta.* 2007;1773: 8–16. doi:10.1016/j.bbamcr.2006.07.005
143. Daugherty RL, Gottardi CJ. Phospho-regulation of  $\beta$ -Catenin Adhesion and Signaling Functions. *Physiology.* 2008;3: 303–309. doi:10.1152/physiol.00020.2007
144. Takahashi K, Sumarriva K, Kim R, Jiang R, Brantley-Sieders DM, Chen J, et al. Determination of the CD148-Interacting Region in Thrombospondin-1. Roberts DD, editor. *PLoS One.* Public Library of Science; 2016;11: e0154916. doi:10.1371/journal.pone.0154916
145. Burrridge K, Sastry SK, Sallee JL. Regulation of cell adhesion by protein-tyrosine phosphatases: I. Cell-matrix adhesion. *J Biol Chem.* 2006;281: 15593–15596. doi:10.1074/jbc.R500030200
146. Sallee JL, Wittchen ES, Burrridge K. Regulation of cell adhesion by protein-tyrosine phosphatases II. Cell-cell adhesion. *J Biol Chem.* 2006;281: 16189–16192. doi:10.1074/jbc.R600003200

147. Bolós V, Peinado H, Pérez-Moreno MA, Fraga MF, Esteller M, Cano A. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci.* 2003;116: 499–511. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12508111>
148. Medici D, Hay ED, Olsen BR. Snail and Slug Promote Epithelial-Mesenchymal Transition through  $\beta$ -Catenin-T-Cell Factor-4-dependent Expression of Transforming Growth Factor- $\beta$ . *Mol Biol Cell.* 2008;19: 4875–4887. doi:10.1091/mbc.E08-05-0506
149. Schneider MR, Wolf E. The epidermal growth factor receptor ligands at a glance. *J Cell Physiol.* 2009;218: 460–466. doi:10.1002/jcp.21635
150. Nakamura T, Sakai K, Nakamura T, Matsumoto K. Hepatocyte growth factor twenty years on: Much more than a growth factor. *J Gastroenterol Hepatol.* 2011;26: 188–202. doi:10.1111/j.1440-1746.2010.06549.x
151. Alcaraz A, Mrowiec A, Insausti CL, García-Vizcaíno EM, Ruiz-Canada C, López-Martínez MC, et al. Autocrine TGF- $\beta$  Induces Epithelial to Mesenchymal Transition in Human Amniotic Epithelial Cells. *Cell Transplant.* 2013;22: 1351–1367. doi:10.3727/096368912X657387
152. Fang D, Hawke D, Zheng Y, Xia Y, Meisenhelder J, Nika H, et al. Phosphorylation of  $\beta$ -Catenin by AKT Promotes  $\beta$ -Catenin Transcriptional Activity. *J Biol Chem.* 2007;282: 11221–11229. doi:10.1074/jbc.M611871200
153. Brembeck FH, Schwarz-Romond T, Bakkers J, Wilhelm S, Hammerschmidt M, Birchmeier W. Essential role of BCL9-2 in the switch between  $\beta$ -catenin's adhesive and transcriptional functions. *Genes Dev.* 2004;18: 2225–2230. doi:10.1101/gad.317604
154. Mierke CT, Frey B, Fellner M, Herrmann M, Fabry B. Integrin  $\alpha 5 \beta 1$  facilitates cancer cell invasion through enhanced contractile forces. *J Cell Sci. Company of Biologists;* 2011;124: 369–83. doi:10.1242/jcs.071985
155. Brunton VG, Avizienyte E, Fincham VJ, Serrels B, Metcalf CA, Sawyer TK, et al. Identification of Src-Specific Phosphorylation Site on Focal Adhesion Kinase: Dissection of the Role of Src SH2 and Catalytic Functions and Their Consequences for Tumor Cell Behavior. *Cancer Res.* 2005;65: 1335–1342. doi:10.1158/0008-5472.CAN-04-1949
156. Westhoff MA, Serrels B, Fincham VJ, Frame MC, Carragher NO. SRC-mediated phosphorylation of focal adhesion kinase couples actin and adhesion dynamics to survival signaling. *J cell Biol.* 2004;24: 8113–33. doi:10.1128/MCB.24.18.8113-8133.2004
157. Nayal A, Webb DJ, Brown CM, Schaefer EM, Vicente-Manzanares M, Horwitz AR. Paxillin phosphorylation at Ser273 localizes a GIT1/PIX/PAK complex and regulates adhesion and protrusion dynamics. *J Cell Biol.* 2006;173: 587–589. doi:10.1083/jcb.200509075
158. Chan PM, Lim L, Manser E. PAK Is Regulated by PI3K, PIX, CDC42, and PP2C and Mediates Focal Adhesion Turnover in the Hyperosmotic Stress-induced p38 Pathway. *J Biol Chem.* 2008;283: 24949–24961. doi:10.1074/jbc.M801728200

159. Li L, Okura M, Imamoto A. Focal adhesions require catalytic activity of Src family kinases to mediate integrin-matrix adhesion. *Mol Cell Biol.* 2002;22: 1203–17. Available: <http://www.ncbi.nlm.nih.gov/pubmed/11809811>
160. Ungewiss C, Rizvi ZH, Roybal JD, Peng DH, Gold KA, Shin D-H, et al. The microRNA-200/Zeb1 axis regulates ECM-dependent  $\beta$ 1-integrin/FAK signaling, cancer cell invasion and metastasis through CRKL. *Sci Rep.* 2016;6: 18652. doi:10.1038/srep18652
161. Chen T-H, Chan P-C, Chen C-L, Chen H-C. Phosphorylation of focal adhesion kinase on tyrosine 194 by Met leads to its activation through relief of autoinhibition. *Oncogene.* 2011;30: 153–166. doi:10.1038/onc.2010.398
162. Peroukides S, Bravou V, Varakis J, Alexopoulos A, Kalofonos H, Papadaki H. ILK overexpression in human hepatocellular carcinoma and liver cirrhosis correlates with activation of Akt. *Oncol Rep.* 2008;20: 1337–44. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19020711>
163. Vi L, de Lasa C, DiGuglielmo GM, Dagnino L. Integrin-Linked Kinase Is Required for TGF- $\beta$ 1 Induction of Dermal Myofibroblast Differentiation. *J Invest Dermatol.* 2011;131: 586–593. doi:10.1038/jid.2010.362
164. Solaini G, Baracca A, Lenaz G, Sgarbi G. Hypoxia and mitochondrial oxidative metabolism. *Biochim Biophys Acta - Bioenerg.* 2010;1797: 1171–1177. doi:10.1016/j.bbabi.2010.02.011
165. Chen J, Imanaka N, Chen J, Griffin JD. Hypoxia potentiates Notch signaling in breast cancer leading to decreased E-cadherin expression and increased cell migration and invasion. *Br J Cancer.* 2010;102: 351–60. doi:10.1038/sj.bjc.6605486
166. Sahlgren C, Gustafsson M V., Jin S, Poellinger L, Lendahl U. Notch signaling mediates hypoxia-induced tumor cell migration and invasion. *Proc Natl Acad Sci.* 2008;105: 6392–6397. doi:10.1073/pnas.0802047105
167. Liu J, Chen B, Lu Y, Guan Y, Chen F. JNK-dependent Stat3 phosphorylation contributes to Akt activation in response to arsenic exposure. *Toxicol Sci.* 2012;129: 363–71. doi:10.1093/toxsci/kfs199
168. Kim J-H, Lee SC, Ro J, Kang HS, Kim HS, Yoon S. Jnk signaling pathway-mediated regulation of Stat3 activation is linked to the development of doxorubicin resistance in cancer cell lines. *Biochem Pharmacol.* 2010;79: 373–80. doi:10.1016/j.bcp.2009.09.008
169. Lee M-H, Koria P, Qu J, Andreadis ST. JNK phosphorylates beta-catenin and regulates adherens junctions. *FASEB J.* 2009;23: 3874–83. doi:10.1096/fj.08-117804
170. Santibañez JF. JNK mediates TGF-beta1-induced epithelial mesenchymal transdifferentiation of mouse transformed keratinocytes. *FEBS Lett.* 2006;580: 5385–5391. doi:10.1016/j.febslet.2006.09.003
171. Zhan X, Feng X, Kong Y, Chen Y, Tan W. JNK signaling maintains the mesenchymal properties of multi-drug resistant human epidermoid carcinoma KB cells through snail and twist1. *BMC Cancer.* 2013;13: 180. doi:10.1186/1471-2407-13-180

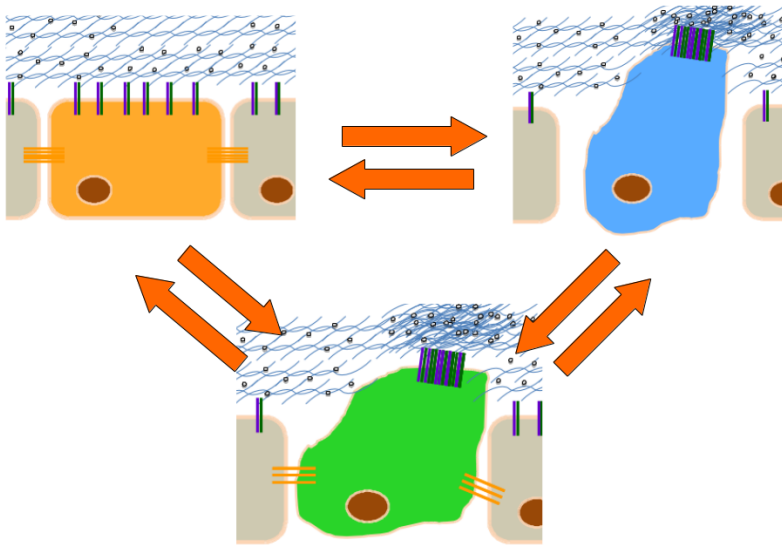
172. Xie G, Yao Q, Liu Y, Du S, Liu A, Guo Z, et al. IL-6-induced epithelial-mesenchymal transition promotes the generation of breast cancer stem-like cells analogous to mammosphere cultures. *Int J Oncol.* 2012;40: 1171–9. doi:10.3892/ijo.2011.1275
173. Bao Y, Hata Y, Ikeda M, Withanage K. Mammalian Hippo pathway: from development to cancer and beyond. *J Biochem.* 2011;149: 361–79. doi:10.1093/jb/mvr021
174. Vasioukhin V, Bauer C, Yin M, Fuchs E. Directed actin polymerization is the driving force for epithelial cell-cell adhesion. *Cell.* 2000;100: 209–19. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10660044>
175. Gumbiner BM. Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol.* 2005;6: 622–34. doi:10.1038/nrm1699
176. Balzac F, Avolio M, Degani S, Kaverina I, Torti M, Silengo L, et al. E-cadherin endocytosis regulates the activity of Rap1: a traffic light GTPase at the crossroads between cadherin and integrin function. *J Cell Sci.* 2005;118: 4765–83. doi:10.1242/jcs.02584
177. Ireton RCRC, Davis MA, van Hengel J, Mariner DJ, Barnes K, Thoreson MA, et al. A novel role for p120 catenin in E-cadherin function. *J cell Biol.* 2002;159: 465–476. doi:10.1083/jcb.200205115
178. Berx G, Staes K, van Hengel J, Molemans F, Bussemakers MJ, van Bokhoven A, et al. Cloning and characterization of the human invasion suppressor gene E-cadherin (CDH1). *Genomics.* 1995;26: 281–9. Available: <http://www.ncbi.nlm.nih.gov/pubmed/7601454>
179. Kim S-HSHSSH, Kim S-HSHSSH. Antagonistic effect of EGF on FAK phosphorylation/dephosphorylation in a cell. *Cell Biochem Funct.* 2008;26: 539–547. doi:10.1002/cbf.1457
180. Choma DP, Milano V, Pumiglia KM, Michael DiPersio C, DiPersio CM. Integrin alpha3beta1-Dependent Activation of FAK/Src Regulates Rac1-Mediated Keratinocyte Polarization on Laminin-5. *J Invest Dermatol.* 2007;127: 31–40. doi:10.1038/sj.jid.5700505
181. Rahimi N, Hung W, Tremblay E, Saulnier R, Elliott B. c-Src kinase activity is required for hepatocyte growth factor-induced motility and anchorage-independent growth of mammary carcinoma cells. *J Biol Chem.* 1998;273: 33714–21. Available: <http://www.ncbi.nlm.nih.gov/pubmed/9837958>
182. Peiro S, Escrivà M, Puig I, Barberà MJ, Dave N, Herranz N, et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res.* 2006;34: 2077–2084. doi:10.1093/nar/gkl141
183. Lemieux E, Bergeron SS, Durand VV, Asselin C, Saucier C, Rivard N. Constitutively active MEK1 is sufficient to induce epithelial-to-mesenchymal transition in intestinal epithelial cells and to promote tumor invasion and metastasis. *Int J Cancer.* 2009;125: 1575–1586. doi:10.1002/ijc.24485
184. de Herreros AG, Peiró S, Nassour M, Savagner P. Snail Family Regulation and Epithelial Mesenchymal Transitions in Breast Cancer Progression. *J Mammary Gland Biol Neoplasia.* 2010;15: 135–147. doi:10.1007/s10911-010-9179-8

185. Thuault S, Tan E-J, Peinado H, Cano A, Heldin C-H, Moustakas A. HMGA2 and Smads Co-regulate SNAIL1 Expression during Induction of Epithelial-to-Mesenchymal Transition. *J Biol Chem*. 2008;283: 33437–33446. doi:10.1074/jbc.M802016200
186. Conacci-Sorrell M, Simcha I, Ben-Yedidia T, Blechman J, Savagner P, Ben-Ze'ev A. Autoregulation of E-cadherin expression by cadherin-cadherin interactions. *J Cell Biol*. 2003;163: 847–857. doi:10.1083/jcb.200308162
187. Unno J, Satoh K, Hirota M, Kanno A, Hamada S, Ito H, et al. LIV-1 enhances the aggressive phenotype through the induction of epithelial to mesenchymal transition in human pancreatic carcinoma cells. *Int J Oncol*. 2009;35: 813–21. Available: <http://www.ncbi.nlm.nih.gov/pubmed/19724917>
188. Yang Z, Rayala S, Nguyen D, Vadlamudi RK, Chen S, Kumar R. Pak1 phosphorylation of snail, a master regulator of epithelial-to-mesenchyme transition, modulates snail's subcellular localization and functions. *Cancer Res*. 2005;65: 3179–84. doi:10.1158/0008-5472.CAN-04-3480
189. Liu Y, Dong Q-Z, Wang S, Xu H-T, Miao Y, Wang L, et al. Kaiso Interacts with p120-Catenin to Regulate beta-Catenin Expression at the Transcriptional Level. *Anto RJ, editor. PLoS One*. 2014;9: e87537. doi:10.1371/journal.pone.0087537
190. Herynk MH, Tsan R, Radinsky R, Gallick GE. Activation of c-Met in colorectal carcinoma cells leads to constitutive association of tyrosine-phosphorylated beta-catenin. *Clin Exp Metastasis*. 2003;20: 291–300. Available: <http://www.ncbi.nlm.nih.gov/pubmed/12856716>
191. Collu GM, Hidalgo-Sastre A, Acar A, Bayston L, Gildea C, Leverenz MK, et al. Dishevelled limits Notch signalling through inhibition of CSL. *Development*. 2012;139: 4405–4415. doi:10.1242/dev.081885
192. Espinosa L, Cathelin S, D'Altri T, Trimarchi T, Statnikov A, Guiu J, et al. The Notch/Hes1 Pathway Sustains NF- $\kappa$ B Activation through CYLD Repression in T Cell Leukemia. *Cancer Cell*. 2010;18: 268–281. doi:10.1016/j.ccr.2010.08.006
193. Stamos JL, Weis WI. The beta-catenin destruction complex. *Cold Spring Harb Perspect Biol*. 2013;5. doi:10.1101/cshperspect.a007898
194. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci*. 2010;123: 4195–4200. doi:10.1242/jcs.023820
195. Stewart RL, O'Connor KL. Clinical significance of the integrin  $\alpha 6 \beta 4$  in human malignancies. *Lab Invest*. 2015;95: 976–986. doi:10.1038/labinvest.2015.82
196. Humphries B, Yang C. The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. *Oncotarget*. 2015;6: 6472–6498. doi:10.18632/oncotarget.3052
197. Krafts KP. Tissue repair: The hidden drama. *Organogenesis*. Taylor & Francis; 2010;6: 225–33. doi:10.4161/org.6.4.12555

# CHAPTER 3

---

## Discussion & Conclusion





## **1. Discussion**

Advances in the study of EMT in cancer context have been limited by its complex regulation involving multiple signals from the tumour microenvironment and multiple intertwined signalling pathways [1,2]. This has been a challenge for systems biology approaches to evaluate multiple hypotheses and generate new testable predictions [3,4]. Hopefully, to elucidate the mechanisms of metastasis and identify new drug targets that could help the design of therapeutic strategies to fight cancer [4].

In this thesis, we developed the first computational model for the regulation of cell adhesion properties involved in EMT and cancer invasion, accounting for multiple microenvironment signals. The model also accounted for main signalling pathways involved in EMT and allows to distinguish single (Mesenchymal phenotype) from collective (Hybrid phenotype) forms of cancer invasion. We analysed the conditions that result in the switching between Epithelial, Mesenchymal and Hybrid phenotypes and accessed the hypothetical role of neighbouring cells signals through the cell-cell contact activation of RPTPs and FAT4. The model analysis revealed that these signals play a role in the control EMT and pinpointed mutational and microenvironment deregulations that favour invasion of carcinomas. In this chapter, we discussed the main advances from this work and their limitations highlighted.

### **1.1. A computational model of cell adhesion**

Within this thesis, we proposed a new computational model for the regulation of cell adhesion properties involved in EMT and carcinomas invasion, based on the logical framework initially proposed by Rene Thomas [5]. Our modelling approach contributed to access multiple combinatorial effects of signals from the tumour microenvironment ( $2^8$  combinations of signals) making this the largest set of

signals analysed so far. In addition, we applied the model to evaluate the impact of single and double mutations in cancer. With this framework, we overcome the lack of kinetic characterization of most processes involved in EMT and provide a simple alternative way that consists on a qualitative view of EMT regulation. In this model, we abstracted the main biochemical regulatory mechanisms involved in EMT into influence effects which constitutes a simplified view of biological processes. Time and dose-dependent effects (e.g. concentration of active molecules) are highly abstracted in this approach, which came with the cost of losing predictive resolution, resulting in a superficial view of the EMT regulation (binary effects). The size of the model variables (51 nodes) also brings noise to predictions, which requires a carefully interpretation of the outcomes. Thus, the results from the model should be interpreted as extreme cases in biology, where the activation of signalling cascades are saturated by the microenvironment signals (strong signal) or basal (having no effect). Since we do not know from the model the exact concentrations or activity thresholds, the predictions from the model may not be verified if the degree of signalling activity in the experimental conditions are between strong and basal.

To make biologically relevant predictions for generic Epithelial cell types, we developed the model based on literature knowledge for the regulation of cell adhesion coming from the curated KEGG PATHWAYS, Atlas of Cancer Signalling Network, together with hundreds of reports of individual effects in Epithelial cell lines [6,7]. Despite of accounting for a considerable amount of reported regulatory effects, in cancer research the literature is too extensive to cover all, with new publications every month. For the microenvironment, we just accounted for a few relevant signals (8 signals) in comparison to the ones a cancer cell can be exposed, particularly in the case of cell-cell contact signalling [8–10]. Thus our model describes only a part of the microenvironment and its predictions should

be taken with reserve, assuming no other signals interfere. Nevertheless, this is the most detailed model in terms of microenvironment signals and accounts for the regulatory processes within the most detailed published EMT network model [11]. Importantly, we have included additional details not yet analysed in the EMT context, such as the integrin signalling, phospho-regulation of the E-cadherin/catenins complex and the Hippo signalling [11,12]. These features have been considered to be quite relevant for the regulation of cell adhesion and EMT signalling pathways to be neglected [13–17]. With these features, we complemented the existent models, challenged our model predictions, and provided new original predictions for the study of EMT in carcinomas. Despite all limitations, the model was successful in recapitulating several biological observations in terms of phenotypes, molecular activity and reported microenvironmental and mutational effects (sections 3.3-3.5 in chapter 2). In principle, the computational model can be applied to any type of Epithelial cells since it was developed without any cell type specificity. Moreover, we recapitulated features from other cellular processes such as the border cells migration and experiments on different Epithelial cell types with and/without cancer. Thus, we developed a model that can also be used in biological problems in other cellular contexts such as development and wound healing.

## **1.2. The role of cell-cell dependent RPTPs**

In this thesis, we found that the activation of RPTPs in our network model by an external signal from neighbouring cells (cell-cell contacts) was capable of keeping the Epithelial adhesion in the presence of growth factors and ECM stiffness (sections 3.6, 3.7, and 3.9 of chapter 2). Importantly, the model predicted that RPTP activation was also capable of switching the phenotypes (Mesenchymal or Hybrid) in favour of Epithelial phenotype (section 3.9 of chapter 2). Although this

is not proven, our prediction already makes an important contribution to the field with the identification of a novel mechanism for the control of EMT [1,2,4]. Moreover, the predicted effect of RPTP activation provides the first mechanistic explanation for the microenvironment induced MET in normal and cancer contexts [18,19]. In this thesis, we propose that a strong activation of RPTP by neighbouring cells would be able to control of EMT. This implies that RPTP activation keeps RTK inhibited even in the presence of RTK activating signals (growth factors). Indeed this is possible based on the 1000-fold higher rate constants of activity of RPTPs in comparison with the ones reported for RTKs [20]. It is expected that this strong activation of RPTPs is stimulated in particular high cell confluence conditions that guarantee enough cell-cell contacts to ensure high stability of RPTP in the membrane, where their targets are localised (catenins and RTKs) [21]. In theory, this implies that EMT is favoured in low numbers of neighbouring cells and MET in high numbers of neighbouring cells, assuming they express RPTPs and their ligands in all cell types (Epithelial, Hybrid and Mesenchymal cells). This would effectively control EMT based on the demand and explain its transient behaviour in wound healing [18]. In cancer, EMT is considered to promote invasion, thus it is expected that cancer cells would overcome this control mechanism by either mutations or inhibitory mechanisms. Several RPTPs have been found to be expressed in most Epithelial tissues and a few have been found to be down-regulated in cancer [21,22]. Thus, our prediction can be applied to an entire class of proteins that requires a join activation, or in alternative a particular type of RPTPs. Only a very small set of RPTPs have been proven to be controlled by cell confluency through cell-cell contacts by homophilic (RPTP-k) or other type of ligands (DEP-1) [21]. Thus our predictions place RPTP-k and DEP1 as been hypothetically involved in the control of EMT that potentially can prevent tumour cells from invading the primary site. This may help in the

future the design of new therapeutic strategies that explore the RPTPs effects, or a particular RPTP activation to prevent tumour cells to invasion primary site via EMT (formation of Mesenchymal tumour cells).

### **1.3. Critical deregulations promoting EMT in carcinomas**

From multiple hypotheses, with our approach we identify few critical deregulations that potentially in abnormal EMT in carcinomas (sections 3.6, 3.7, and 3.8 of chapter 2). The tumour microenvironment composed of hypoxia together with signals from chronic inflammation was shown to result in an irreversible EMT behaviour (section 3.8 in chapter 2). SRC overexpression was also shown to be result in a irreversible EMT behaviour (section 3.10 of chapter 2). Such irreversible behaviour would lead to the accumulation of Mesenchymal cells in the tumour, placing these perturbations to potentially favour invasion of carcinomas. Both deregulations are found in carcinomas and correlated with the metastable phenotype [23,24]. Further, these behaviours were associated to the capacity of these deregulations to inhibit RPTPs activity by either an oxidation mechanism (hypoxia) or an transcriptional mechanism (SRC). Once these perturbations inhibit RPTPs, EMT no longer is able to be controlled by our proposed mechanism. This places Hypoxia and SRC overexpression in chronic inflammation microenvironment as deregulations in cancer that are enough to promote invasion through Mesenchymal cells. Thus, our modelling work provides a new mechanistic explanation for the acquisition of invasion properties in carcinomas through EMT. These predictions are consistent with the idea of both mutational or microenvironment origin of metastasis, and supports the idea of SRC inhibitors and antioxidants usage as therapeutic strategies to prevent metastasis. However, our predictions are conditioned by the supporting data associated to RPTP-k, which could not be true for other RPTP such as DEP-1.

This means that other RPTPs could be insensitive to ROS or not subject to SRC mediated transcriptional regulation. Nevertheless, because RPTP-k is always expressed in Epithelial cells it is expected at least the observation of some mild effects experimentally [21].

#### **1.4. The stability of Hybrid phenotypes**

Our modelling work explored both microenvironment and mutational causes of the stability of Hybrid phenotypes, which are implicated in ensuring collective cell migration of carcinomas (sections 3.7 of chapter 2). Our workflow provided the identification of some putative perturbations that could promote the stability of hybrid phenotypes. Some are clinically relevant such as the cases of SMAD and TGF- $\beta$  inhibitions, which provides an warning for the usage of drugs to inhibit this targets in cancer therapy. Strikingly, we also found that the uncontrolled expression (overexpression) of FAT4 in cancer have the potential to stabilize the Hybrid adhesion phenotype. Our finding provided a putative novel role for the human FAT4 genes, frequently overexpressed in carcinomas. These genes are usually considered as tumour suppressors by their capacity to result in tumour growth inhibition, making this finding unexpected [25]. Our modelling work, further predicts that the stability of the Hybrid phenotype requires the presence of microenvironment signals that are considered as drivers of EMT (HGF, EGF and ECM stiffness). This makes sense since Hybrid phenotypes result from partial EMT and also explains how SMAD and TGF- $\beta$  inhibitions can lead to Hybrid phenotype. Interestingly, Hybrid stability is predicted to depend on cell-cell contact signals from neighbouring cells for the activation of FAT4 by its ligands without activating RPTP. In principle, this could be achieved either by hypoxia, SRC overexpression or by mutations in RPTP ligands. Hypoxia is the most plausible, which could explain the switching behaviour when tumour cells

colonize normal tissues. Although Hybrid stability is crucial for collective cell migration, the formation of the clusters is expected to be more complex. Further work in a multicellular context is required to understand how these clusters are formed and how invasion and colonization operate. Taken all together, we propose the FAT4 overexpression as a critical alteration that is a promising candidate to promote collective cell migration in carcinomas.

## **2. Conclusions**

Based on the results from the computational model presented in this thesis, the following conclusions can be outlined:

- A novel computational model was developed for generating predictions of adhesion properties characteristic of distinct modes of migration accounting for multiple microenvironment signals, which can be applied to other questions and cellular contexts.
- Cell-cell contact signals that strongly activate RPTPs could prevent EMT and stimulate MET in the presence of growth factors and ECM stiffness, which is novel mechanism to control EMT.
- The tumour microenvironment (hypoxia with chronic inflammation signals) and the reported SRC overexpression in cancer are critical deregulations that would favour carcinomas invasion by overcoming the control of EMT by RPTPs.
- The overexpression of FAT4 may stabilize collective migration of carcinomas in the presence of EMT driving signals and neighbouring cells signals (cell-cell contacts) that only activate FAT4 receptors (not activate RPTPs), a novel function for FAT4 genes.

Taken together, the predictions generated in this thesis make an important contribution to the understanding of the how EMT is controlled by signals in the

tumour microenvironment. In this work, we raised new testable hypotheses that explain invasion of carcinomas and pinpointed critical alterations that favour single and collective invasion. These predictions further highlighted the role of neighbouring cells signals in cancer invasion and metastasis.

### **3. Future work**

In this thesis, we provided predictions that imply novel mechanisms that hypothetically could be used to prevent metastasis and thus they should be experimentally tested in a future work. The most important to be tested is the hypothetical role of RPTP as a control mechanism of EMT and its disruption by SRC overexpression. At the time of the writing of this thesis, *in vitro* experiments that test the effect of RPTP-k overexpression in breast cancer cell lines with and without SRC overexpression/overactivation already started by the Lab of Florence Janody. It would be also relevant to test the hypothesis that the RPTP-k needs to be stabilized in the membrane by cell-cell contacts to have an effect. Additional modelling work that explores the kinetic properties of RPTP-k mechanisms in combination with these experiments would also help to understand in depth how to control EMT in the cancer context.

The predicted role of FAT4 overexpression in driving collective cell migration in our modelling work would also be relevant to test experimentally *in vitro* for carcinoma cell lines. Importantly, studies that better characterize the RPTP and FAT4 ligands in Epithelial, Mesenchymal and Hybrid cell types in terms of expression and regulation would provide additional insights on the role of neighbouring cells in cancer. Extending our model to a multicellular context using cellular automata tools (e.g. Epilog) would provide further insights on the dynamics of how collective invasion and metastasis of tumours is formed under the influence of neighbouring cells signals.



In cancer therapy, the RPTP EMT control mechanism and FAT4 effect on the stability of hybrid phenotype highlighted in this modelling work can also be explored, in particular for testing combinations of drugs that improve the control of cancer spreading. In the case of RPTP, drugs that target its stability in the membrane in combination with antioxidants (to prevent its inactivity by ROS) would be potential candidates to reduce cancer invasion. In addition, additional combination with drugs that target SRC would be a fruitful candidate in the case of cancers that overexpress SRC. In the case of cancer that overexpress FAT4, the development and usage of drugs that destabilize FAT4 in the membrane may additionally help in preventing cancer invasion via the formation of collective cell migration. Furthermore, the predicted effects of inhibiting MAPK signalling and the combination of AKT/MET with CK1 inhibition should also be explored as possible drug targets to prevent cancer invasion. Together, these experiments and additional modelling work would undoubtedly help clarify the fascinating and unexpected roles of the RPTP and FAT4 by cell-cell contacts in cancer.

#### 4. References

1. Gao D, Vahdat LT, Wong S, Chang JC, Mittal V. Microenvironmental regulation of epithelial-mesenchymal transitions in cancer. *Cancer Res. NIH Public Access*; 2012;72: 4883–9. doi:10.1158/0008-5472.CAN-12-1223
2. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol. Nature Publishing Group*; 2014;15: 178–96. doi:10.1038/nrm3758
3. Somvanshi PR, Venkatesh K V. A conceptual review on systems biology in health and diseases: from biological networks to modern therapeutics. *Syst Synth Biol. Springer*; 2014;8: 99–116. doi:10.1007/s11693-013-9125-3
4. Jolly MK, Tripathi SC, Somarelli JA, Hanash SM, Levine H. Epithelial/mesenchymal plasticity: how have quantitative mathematical models helped improve our understanding? *Mol Oncol. Wiley-Blackwell*; 2017;11: 739–754. doi:10.1002/1878-0261.12084
5. Thomas R. Regulatory networks seen as asynchronous automata: A logical description. *J Theor Biol. Academic Press*; 1991;153: 1–23. doi:10.1016/S0022-5193(05)80350-9

6. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res. Oxford University Press*; 2017;45: D353–D361. doi:10.1093/nar/gkw1092
7. Kuperstein I, Bonnet E, Nguyen H-AH, Cohen D, Viara E, Grieco L, et al. Atlas of Cancer Signalling Network: a systems biology resource for integrative analysis of cancer data with Google Maps. *Oncogenesis*. 2015;4: 1–14. doi:10.1038/oncsis.2015.19
8. Deneff C. Contact-dependent Signaling. *Cell Commun Insights*. 2014; 1–11. doi:10.4137/CCI.s12484.TYPE
9. Br ucher BLD, Jamall IS. Cell-Cell Communication in the Tumor Microenvironment, Carcinogenesis, and Anticancer Treatment. *Cell Physiol Biochem*. 2014;34: 213–243. doi:10.1159/000362978
10. Li H, Xu F, Li S, Zhong A, Meng X, Lai M. The tumor microenvironment: An irreplaceable element of tumor budding and epithelial-mesenchymal transition-mediated cancer metastasis. *Cell Adh Migr. Taylor & Francis*; 2016;10: 434–446. doi:10.1080/19336918.2015.1129481
11. Steinway SN, Zanudo JGT, Ding W, Rountree CB, Feith DJ, Loughran TP, et al. Network modeling of TGF $\beta$  signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint Sonic hedgehog and Wnt pathway activation. *Cancer Res*. 2014;74: 5963–5977. doi:10.1158/0008-5472.CAN-14-0225
12. Cohen DPA, Martignetti L, Robine S, Barillot E, Zinovyev A, Calzone L. Mathematical Modelling of Molecular Pathways Enabling Tumour Cell Invasion and Migration. *PLoS Comput Biol*. 2015;11. doi:10.1371/journal.pcbi.1004571
13. Daugherty RL, Gottardi CJ. Phospho-regulation of  $\beta$ -Catenin Adhesion and Signaling Functions. *Physiology*. 2008;3: 303–309. doi:10.1152/physiol.00020.2007
14. Le Bras GF, Taubenslag KJ, Andl CD. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adh Migr. Taylor & Francis*; 2012;6: 365–73. doi:10.4161/cam.21326
15. Varelas X, Miller BW, Sopko R, Song S, Gregorieff A, Fellouse FA, et al. The Hippo pathway regulates Wnt/beta-catenin signaling. *Dev Cell*. 2010;18: 579–91. doi:10.1016/j.devcel.2010.03.007
16. Canel M, Serrels A, Frame MC, Brunton VG. E-cadherin–integrin crosstalk in cancer invasion and metastasis. *J Cell Sci*. 2013;126: 393–401. doi:10.1242/jcs.100115
17. Winograd-Katz SE, F  ssler R, Geiger B, Legate KR. The integrin adhesome: from genes and proteins to human disease. *Nat Rev Mol Cell Biol*. 2014;15: 273–288. doi:10.1038/nrm3769
18. Leopold PL, Vincent J, Wang H. A comparison of epithelial-to-mesenchymal transition and re-epithelialization. *Semin Cancer Biol*. 2012;22: 471–483. doi:10.1016/j.semcancer.2012.07.003

19. Ding S, Zhang W, Xu Z, Xing C, Xie H, Guo H, et al. Induction of an EMT-like transformation and MET in vitro. *J Transl Med. BioMed Central*; 2013;11: 164. doi:10.1186/1479-5876-11-164
20. Hertog J den, Östman A, Böhmer F-D. Protein tyrosine phosphatases: regulatory mechanisms. *FEBS J. Blackwell Publishing Ltd*; 2008;275: 831–847. doi:10.1111/j.1742-4658.2008.06247.x
21. Xu Y, Fisher GJ. Receptor type protein tyrosine phosphatases (RPTPs) - roles in signal transduction and human disease. *J Cell Commun Signal. Springer*; 2012;6: 125–38. doi:10.1007/s12079-012-0171-5
22. Du Y, Grandis JR. Receptor-type protein tyrosine phosphatases in cancer. *Chin J Cancer. BioMed Central*; 2015;34: 61–9. doi:10.5732/cjc.014.10146
23. Irby RB, Yeatman TJ. Role of Src expression and activation in human cancer. *Oncogene*. 2000;19: 5636–5642. Available: <http://www.nature.com/onc/journal/v19/n49/pdf/1203912a.pdf>
24. Gilkes DM, Semenza GL, Wirtz D. Hypoxia and the extracellular matrix: drivers of tumour metastasis. *Nat Rev Cancer*. 2014;14: 430–439. doi:10.1038/nrc3726
25. Katoh M. Function and cancer genomics of FAT family genes (Review). *Int J Oncol*. 2012;41: 1913–8. doi:10.3892/ijo.2012.1669

Oeiras, October, 2018

## How cells initiate Epithelial-to-Mesenchymal Transition

Ricardo J. Pais



ITQB-UNL | Av. da República, 2780-157 Oeiras, Portugal  
Tel (+351) 214 469 100 | Fax (+351) 214 411 277

**[www.itqb.unl.pt](http://www.itqb.unl.pt)**